CONSTRAINT-DRIVEN RF TEST STIMULUS GENERATION AND BUILT-IN TEST

A Dissertation
Presented to
The Academic Faculty

by

Selim Sermet Akbay

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2010

CONSTRAINT-DRIVEN RF TEST STIMULUS GENERATION AND BUILT-IN TEST

Approved by:

Dr. Abhijit Chatterjee, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. David C. Keezer
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Linda S. Milor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Gregory D. Durgin
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Suresh K. Sitaraman
The George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: December 8th, 2009

To my mother Leyla Tunarli,

to my dear wife Nazmiye,

and my beautiful daughter Defne.

Your love means everything to me.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my adviser, Prof. Abhijit Chatterjee, for providing me not only the opportunity to work with him, but also his invaluable guidance for my research. I would also like to extend my gratitude to my defense committee: Prof. David Keezer, Prof. Linda Milor, Prof. Gregory Durgin, and Prof. Suresh Sitaraman. I appreciate their time and effort serving on my committee and offering valuable suggestions and recommendations. I would also like to thank Semiconductor Research Corporation (SRC), GigaScale Research Center (GSRC), National Science Foundation (NSF), and Packaging Research Center (PRC) for the support during various stages of my graduate studies at Georgia Tech. I am especially grateful to TM Mak from Intel Corporation for his excellent guidance throughout my graduate studies, and to Craig Force from Texas Instruments for his valuable insights and feedback on defect-based test approach.

I extend my thanks to all members of Testing and Reliability Group for the friendly atmosphere and professional support. I would especially like to mention Achintya Halder, Donghoon Han, Ganesh Srinivasan, Hyun Choi, Maryam Ashouei, Mudassar Nisar, Raj Senguttuvan, Shalabh Goyal, Shreyas Sen, Soumendu Bhattacharya, Utku Diril, Vishwanath Natarajan, and Yuvraj Dhillon, in alphabetical order. I also would like to acknowledge Jose L. Torres and Julie M. Rumer from Intel Corporation for their constant support on the IP-constraint project, as well as want to thank them for the great team we were. It is their ability to melt professionalism with their warm hearts that landed me at Intel after Georgia Tech.

I would also like to thank my undergraduate research adviser, Prof. Tayfun Akin, who provided curious young minds a hands-on school within a school. It is

his METU-MEMS Lab that fueled my journey through Georgia Tech and Intel Corporation.

Special thanks go to my sweet wife Nazmiye, who was so patient with me, without her faithful support in writing this thesis I could have given up a long time ago. She was always there for me with constant encouragement and help. I also would like to thank my beautiful daughter Defne, who gave me another big motivation to complete this dissertation.

Serving Turkish community was an important and fun part of my life in Atlanta. In return, I made lifelong friends and invaluable memories. I would like to mention Akay Izat and Ayse Topcuoglu, Alper Akyildiz, Alper Eken, Altan Gulcu, Bortecene Terlemez, Bulent Anil, Cagatay Candan, Erdem Matoglu, Eylem and Ozlem Ekici, Faik Baskaya, Fatma Caliskan, Fusun Ercan, Guclu Onaran, Hasan Ertas, Ilkay Yavrucuk, Maz and Winnie Kosma, Meltem Alemdar, Namik and Meral Ciblak, Nevbahar Ertas, Ozgur Kurc, Serap and Jason Alvarez, Sinasi Bilsel and Gonul Ertem, and Utku Diril, in alphabetical order.

I also would like to thank my mother, brother and immediate family: Leyla Tunarli, Serkan Akbay, Sevket Tunarli, Vildan Caglar, Mehmet Caglar, Mustafa Tunarli, and Mehmet Tunarli, for encouraging me to follow my passion in science, and providing support when it was hard to make ends meet for honest people living in Turkey.

Contents

List of Tables

List of Figures

GLOSSARY

1:3MIX mixture of linear and gaussian distributions in a 1:3 sample ratio.

1dBComp 1dB compression point.

AC accuracy.

ACPR adjacent-channel power ratio.

ADC analog-to-digital converter.

ASIC application-specific integrated circuit.

ATE automated test equipment.

ATG automatic test generation.

ATPG automatic test pattern generation.

AWG arbitrary waveform generator.

BER bit error rate.

BIST built-in self-test.

BIT built-In test.

BOST built-off self-test.

BOT built-off test.

CE classification error.

D2S differential-to-single-ended converter.

DAC digital-to-analog converter.

DECT digital enhanced cordless telecommunications.

DfT design-for-test.

DNL differential nonlinearity.

DOE design of experiments.

DOT defect-oriented test.

DSP digital signal processor.

DUT device under test.

F+ false positive.

F- false negative.

FC fault coverage.

FDART frequency domain alternate response "testimator".

FFT fast fourier transform.

FN false negative.

FP false positive.

GA genetic algorithm.

GAbS genetic-algorithm based search algorithm.

GradS gradient-based search algorithm.

HB harmonic balance.

HDI high-density interconnect.

HVM high-volume manufacturing.

IF intermediate frequency.

IIP3 input referred third order intercept point.

IMD intermodulation.

INL integral nonlinearity.

I/O input/output.

IP intellectual property.

ISWR input standing wave ratio.

KGD known good die.

KGEM known good embedded module.

KGES known good embedded substrate.

LARS least angle regression.

LIN linear distribution.

LNA low-noise amplifier.

LO local oscillator.

LPF lowpass filter.

MARS multivariate adaptive regression splines.

MART multiple additive regression trees.

MC monte carlo.

MCM multi-chip module.

MEMS micro-electro mechanical systems.

mpe maximum percentage error.

MTPR missing tone power ratio.

NARD gaussian distribution around RSM corners.

NF noise figure.

NORM gaussian distribution.

OBT oscillation-based test.

OIP3 output referred third-order intercept point.

OSWR output standing wave ratio.

PA power amplifier.

PC pass coverage.

PE prediction error.

PLL phase-locked loop.

POBT predictive OBT.

PSD power spectral density.

PSS periodic steady state.

RB relay board.

RDB relay driver board.

RevIso reverse isolation.

RF radio frequency.

rms root-mean-square.

rmspe root-mean square percentage error.

RndW random walk search algorithm.

RSM response surface model.

RSMD linear distribution around RSM corners.

SFDR spurious-free dynamic range.

SIB socket interface board.

SMO sequential minimal optimization.

SNR signal-to-noise ratio.

SoC system-on-chip.

SOP system-on-a-package.

SPOT specification-oriented test.

SUT specification under test.

SVM support vector machine.

T+ true positive.

T- true negative.

TCK test clock.

TD time domain.

TDI test data in.

TDO test data out.

TMS test mode select.

TN true negative.

TP true positive.

TRST test reset.

VCO voltage controlled oscillator.

SUMMARY

With the explosive growth in wireless applications, the last decade witnessed an ever-increasing test challenge for radio frequency (RF) circuits. While the design community has pushed the envelope far into the future, by expanding CMOS process to be used with high-frequency wireless devices, test methodology has not advanced at the same pace. Consequently, testing such devices has become a major bottleneck in high-volume production, further driven by the growing need for tighter quality control [1].

RF devices undergo testing during the prototype phase and during high-volume manufacturing (HVM). The benchtop test equipment used throughout prototyping is very precise yet specialized for a subset of functionalities. HVM calls for a different kind of test paradigm that emphasizes throughput and sufficiency, during which the projected performance parameters are measured one by one for each device by automated test equipment (ATE) and compared against defined limits called specifications. The set of tests required for each product differs greatly in terms of the equipment required and the time taken to test individual devices. Together with signal integrity, precision, and repeatability concerns, the initial cost of RF ATE is prohibitively high. As more functionality and protocols are integrated into a single RF device, the required number of specifications to be tested also increases, adding to the overall cost of testing, both in terms of the initial and recurring operating costs.

In addition to the cost problem, RF testing proposes another challenge when these components are integrated into package-level system solutions. In systems-on-packages (SOPs), the test problems resulting from signal integrity, input/output (I/O) bandwidth, and limited controllability and observability have initiated a paradigm shift in

high-speed analog testing, favoring alternative approaches such as built-in tests (BITs) where the test functionality is brought into the package. This scheme can make use of a low-cost external tester connected through a low-bandwidth link in order to perform demanding response evaluations, as well as make use of the analog-to-digital converters (ADCs) and the digital signal processors (DSPs) available in the package to facilitate testing. Although research on analog BIT has demonstrated hardware solutions for single specifications, the paradigm shift calls for a rather general approach in which a single methodology can be applied across different devices, and multiple specifications can be verified through a single test hardware unit, minimizing the area overhead.

Specification-based alternate test methodology provides a suitable and flexible platform for handling the challenges addressed above. In this thesis, a framework that integrates ATE and system constraints into test stimulus generation and test response extraction is presented for the efficient production testing of high-performance RF devices using specification-based alternate tests. As illustrated in Figure 1, the main components of the presented framework are as follows:

- Constraint-driven RF alternate test stimulus generation: An automated test stimulus generation algorithm for RF devices that are evaluated by a specification-based alternate test solution is developed. The high-level models of the test signal path define constraints in the search space of the optimized test stimulus. These models are generated in enough detail such that they inherently define limitations of the low-cost ATE and the I/O restrictions of the device under test (DUT), yet they are simple enough that the non-linear optimization problem can be solved empirically in a reasonable amount of time.

- Feature extractors for BIT: A methodology for the built-in testing of RF devices integrated into SOPs is developed using additional hardware components.

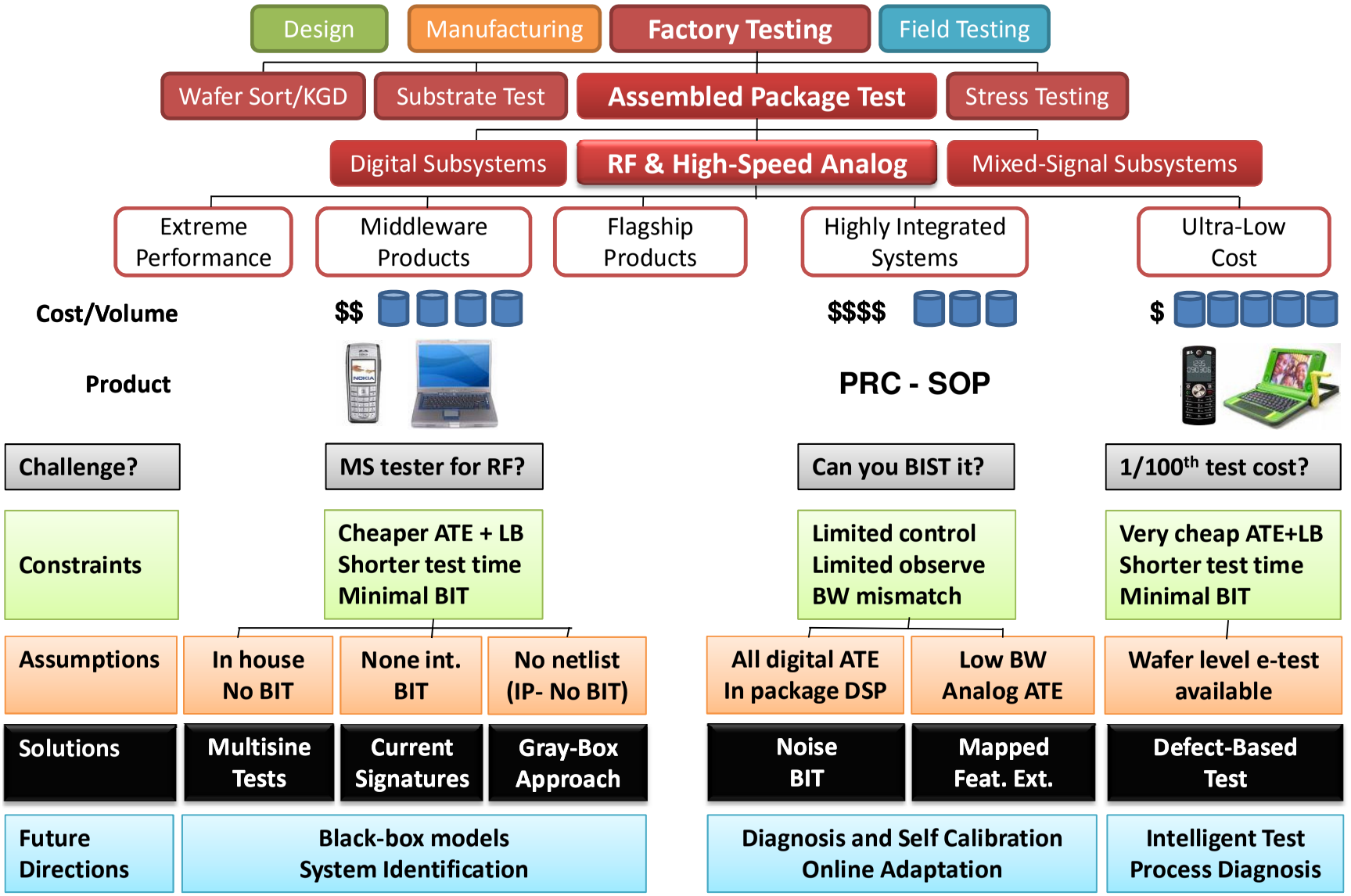


Figure 1: Proposed framework for constraint-driven RF alternate test stimulus generation, BIT with feature extractors, and defect-based tests.

These hardware components correlate the high-bandwidth test response to low bandwidth signatures while extracting the test-critical features of the DUT. Supervised learning is used to map these extracted features, which otherwise are too complicated to decipher by plain mathematical analysis, into the specifications under test (SUT).

- Defect-based alternate testing of RF circuits: A methodology for the efficient testing of RF devices with low-cost defect-based alternate tests is developed. The signature of the DUT is probabilistically compared with a class of defect-free device signatures to explore possible corners under acceptable levels of process parameter variations. Such a defect filter applies discrimination rules generated by a supervised classifier and eliminates the need for a library of possible catastrophic defects.

Chapter I

INTRODUCTION

The life cycle of an RF component starts with a design concept followed by simulations. Once the design is verified, a prototype is fabricated. The first step of testing is conducted on a few prototype samples to ensure the correctness of the design and viability of the manufacturing process. The design verification procedure consists of the application of a set of verification tests followed by design diagnosis. If necessary, the design can be altered, new prototypes can be fabricated, and a new set of tests can be applied until the design is finalized [2, 3], as shown in Figure 2.

Following design debugging, characterization tests are performed, which extensively test all critical design parameters of the final design and make sure that all the performance goals are met. Characterization tests are performed on a sample set of devices, and they also document the variations in device performance. Both verification and characterization tests make use of benchtop test equipment, which is a collection of test devices specialized for different measurements. These tests emphasize completeness and accuracy rather than testing time. Characterization tests, in addition, include functional tests with known good dice attached. Such tests emulate typical operations of the modules and check for the integrity of the overall system. Functional tests require a significant amount of time, from scores of minutes to hours, because many combinations of possible modes of operation need to be validated.

Once an economically feasible yield is attained, the production line goes into HVM. However tightly controlled the process, the engineering tolerances of machines, methodologies, and environmental control define variations in a semiconductor manufacturing line, which affect the electrical properties of individual devices. Because of

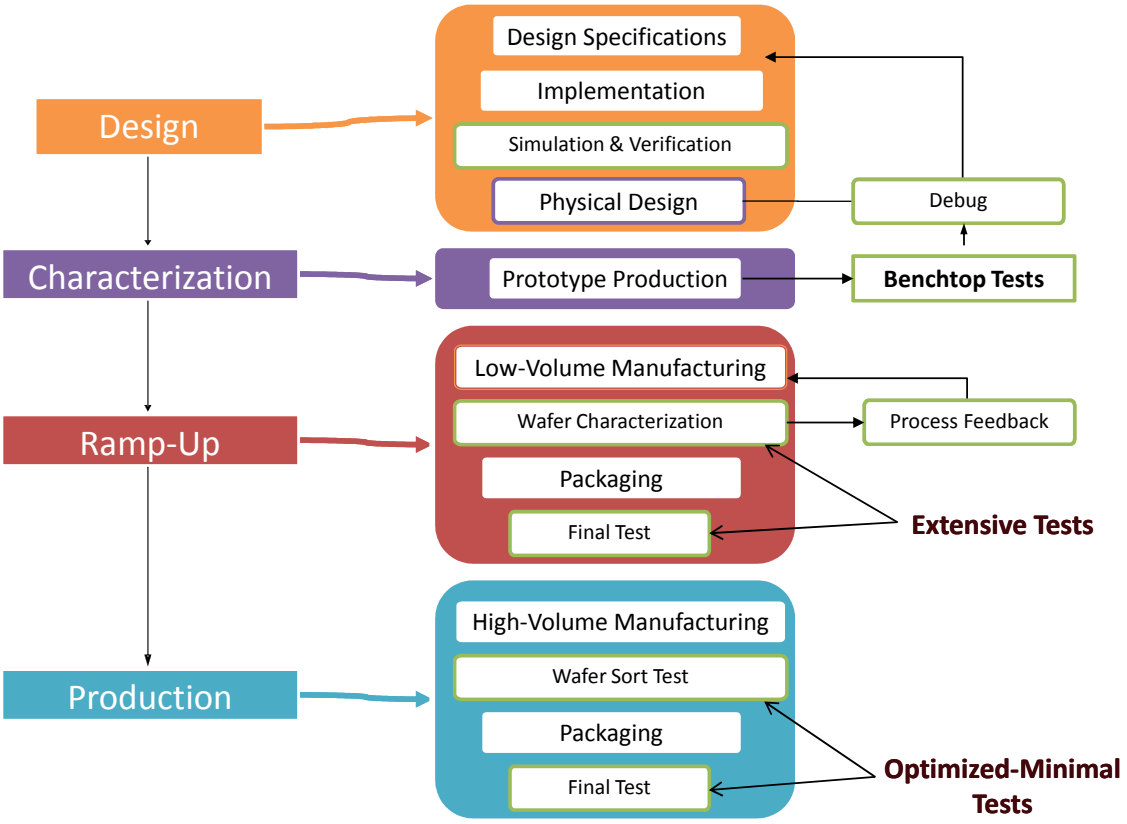


Figure 2: Manufacturing steps for an RF component.

the small sizes of electrical components, there is still a finite probability that any one product may fail to perform at the specified level. These performance benchmarks are called specifications of the device and are listed on the datasheet. The manufacturer carefully tests each device against predefined pass-fail thresholds before the device can be shipped to the customer.

1.1 Constraint-Driven RF Test Stimulus Generation

HVM is performed in two phases: wafer-level test and final test. A wafer-level test checks the generic characteristics of a fabricated wafer to catch catastrophic process problems. It also checks a limited number of specifications of each die to prevent defective dice from being passed through further production steps. A final test is performed on the packaged dice.

HVM calls for a test paradigm that emphasizes throughput and sufficiency. These tests are performed for every one of the units shipped to customers; hence, functional tests cannot be afforded at this level. Also, HVM tests are performed by an ATE rather than a collection of benchtop equipment. The ATE integrates a subset of the functionality covered by the characterization test, yet it is a generic device that can be programmed to use different test plans for different DUTs. The HVM test is fully automated, as shown in Figure 3. First, a handler places one or a few manufactured devices onto a loadboard with sockets for each device. The socket provides a temporary connection to the signal paths on the loadboard, which connects different ATE ports to the DUTs. ATE provides the control signals and test stimuli for many sequential tests and collects the corresponding responses. Each response is compared to a decision boundary. Then ATE evaluates the results for all decision boundaries and assigns the DUT to an appropriate bin. A simple bin structure has pass and fail classes, while devices that are more complicated may be sorted into multiple bins, depending on their performance. Once the proper bin is determined, the handler breaks the mechanical connection and moves the DUT into the proper pile. All these events have to be performed in a very short amount of time given that all manufactured devices need to pass through HVM testing. Depending on the complexity of the device, the test time varies from a fraction of a second to a few seconds; hence, it is one of the more significant factors defining the manufacturing cost per device.

Conventionally, production testing of RF devices is performed in a sequential manner mimicking individual characterization tests. Each step covers a single specification of the device: (i) the specific instruments on the ATE are selected and set up, (ii) the selected instruments are routed to the appropriate input and output ports, possibly using mechanical relays on the loadboard, (iii) the particular test stimulus is applied, and (iv) the test response is measured. The process continues with the next specification, as shown in Figure 4. As a result, the total testing time is linearly

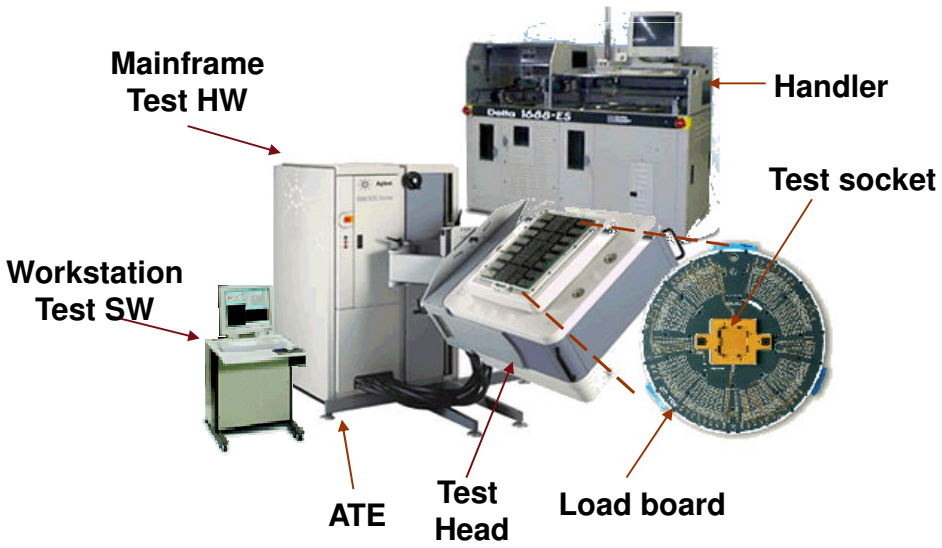


Figure 3: Components of HVM test flow: ATE (pictured Agilent 93000), handler (pictured DELTA 1688), loadboard, and test socket.

proportional to the number of specifications on the datasheet. Thus, testing time increases dramatically as extra functions are integrated into the device. This test time trend proposes a bottleneck for the economics of the integration paradigm in the sense that with smaller feature sizes, more features can be crammed into the same silicon area, and yet the final solution may no longer be economically competitive because the increased testing cost dominates overall production costs. Similarly, extra functions often require additional measurement instruments to be integrated into the ATE, which increases the already prohibitive fixed cost [4].

Several approaches have been proposed to reduce RF test cost, including test selection, test ordering, and fault-based test pattern generation. Test ordering techniques determine an optimal order for sequencing the tests to minimize production testing time. Algorithms are driven to drop tests with low fault coverage, and to determine an optimal subset of tests such that the test time for desired fault coverage is minimized. In contrast, test selection techniques exploit the correlation between circuit performance and a specific test set to reduce the number of specifications that need

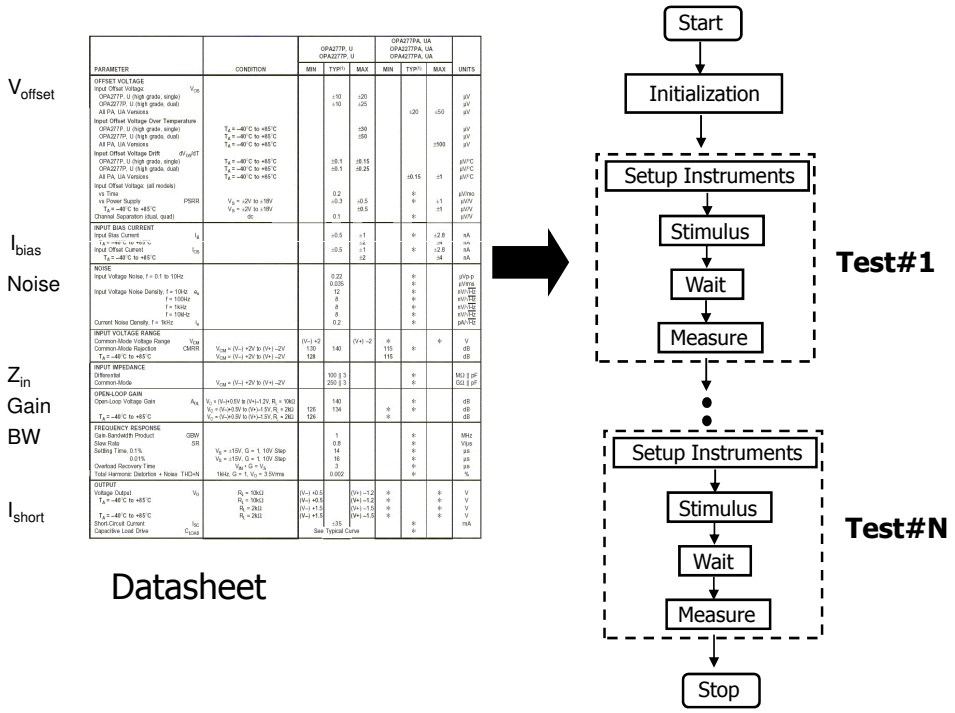


Figure 4: The standard production test model, sequential tests are applied for each specification on the datasheet.

to be explicitly tested. Regression models are constructed to predict the untested performances from the explicitly tested ones. However, both of these techniques focus only on the tests that are present in the original test set. Thus, the test coverage and test time improvements possible through these techniques largely depend on the redundancy and speed of the tests in the original test set. Fault-based test generation techniques generate test stimuli to facilitate the detection of faults in analog circuits. Typically, the faults considered are catastrophic in nature as opposed to parametric, and the tests are designed to be sensitive to these faults. Since there is no explicit relationship between the fault-based tests and the datasheet specifications, this method represents a big shift in the testing paradigm from traditional specification testing. Analog manufacturers normally need to know not only if the device is functioning, but also if the device meets the specifications described in the product datasheet, limiting fault-based test usage in RF production test environments.

Specification-based test generation techniques attempt to bridge the difference between traditional specification testing and fault-based testing by using alternate tests. Testing is based on measurements that are inexpensive to perform compared to conventional specification tests. The pass/fail outcome of a specification test and classification of the DUT into an appropriate performance bin is accomplished by nonlinear regression mapping functions. Since the alternate tests can capture most of the necessary information about a device specification, this testing technique is more in conformance with conventional approaches. In alternate testing, the quality of a test is determined by how much information about the circuit's test specifications is included in the obtained test response. A suite of sequential specification tests is replaced with a single test consisting of a carefully crafted test stimulus applied to the DUT. The response is concurrently mapped to all SUTs by using supervised learning, resulting in significant test time savings. The cost of the external tester and loadboard is also reduced since test stimulus is crafted in a way to favor simpler generators and samplers.

In this thesis, an automatic test generation (ATG) method is presented for alternate testing of RF devices by using cheaper, simpler pattern generators and response analyzers. This simplified ATE provides a scalable and low-cost implementation, and the individual constraints of the ATE can be incorporated into the optimization process as parameters. Several stimulus optimization methods are investigated, and comparisons are documented, including a genetic algorithm (GA) that makes use of the frequency-domain stimulus representation, and with behavioral models to boost the simulation efficiency for DUTs, including third-party intellectual property (IP) blocks. The results show that the presented alternate test stimulus generation approach provides significant test time savings and can be implemented with low-cost testers. Although originally tailored for test operations with full access to netlist and process data, the presented modification can also be applied to IP blocks with

success. The ATE example shows 36% reduction in test time and a 48% reduction in tester cost for a test case handled with behavioral models. Variations in implementation of the ATG methodology are also studied, and results show that there is no one-size-fits-all recipe. The extensive case study in Chapter 3 highlights important factors and their interactions.

1.2 Feature Extractors for Built-In Tests

Scaling has been the fundamental driver for semiconductor manufacturing at the device level. Besides design rules, good process control, and design for test/yield, scaling has been the dominant enabler for manufacturing more complex devices with financially sustainable yields. By employing the same principle, SOP drives system integration just like Moore's law drives IC integration. Testing, like all other components of SOP, needs to be scalable. Otherwise, a larger percentage of manufacturing costs will be allocated to testing, and eventually testing will dominate the overall cost. On the other hand, as functionality increases, more functionality needs to be tested, increasing testing effort and cost. Therefore, HVM testing of SOPs needs to use alternate test methods rather than classical functional tests. In traditional systems using standard surface-mount components, conventional test methods work since the components are minimally tested before being assembled. However, in a highly integrated system such as SOP, new test methods are necessary since the components are integrated into the layers of the package and not all the nodes of the device are accessible.

Manufacturing tests for an SOP involve additional steps. These systems bring together not only analog and digital subsystems but also RF components for high-bandwidth communication, optics as the media for multi-gigahertz data transfer, and micro electro-mechanical systems (MEMS) with moving elements as an interface with the outside world. The design of such systems makes use of many advantages that

are not present in the traditional design flow:

- Ultra-miniaturization and uniformity of the medium in these systems result in shorter data paths and smaller parasitics and thus increase the speed and bandwidth of communication subsystems far beyond classical techniques.
- Using the SOP approach, the designer is no longer limited by the discrete properties of passives but can make use of embedded passive components tailored specifically to the application.
- SOP versions of sensor-processor, actuator-driver, and electrical-optical subsystems benefit from increased bandwidth, minimized package parasitics, and lower power consumption compared to traditional interfaces [5].

Although some components of an SOP can be individually tested at the bare die level, their functionality has to be verified after integration. Furthermore, some components and some specifications are only defined after the integration. As a result, testing can constitute up to 45% of the overall manufacturing cost [5]. These HVM tests are geared toward two main types of failure mechanisms: (i) catastrophic faults, which impair the ability of the device to function altogether, and (ii) parametric faults, which make the device fall short of satisfying the specifications related to a function. Hence, a successful HVM test strategy not only lets out defect-free SOPs, but also makes sure that their performances meet the specifications listed on the datasheet within sufficient margins.

SOP comes with its own challenges in terms of HVM testing. A typical SOP encapsulates many of its internal functions, and production testing is performed by the application of test signals to the SOP using external ATE. The key problem is that the external ATE does not have direct access to all the internal embedded functions of the SOP. It may be possible to route some of the internal electrical signals out of the package to the external tester; however, these internal signals operate at frequencies

that cannot be observed directly by an external tester due to the frequency limitations of the encapsulating package and lower speed of external I/O. A similar speed and integrity concern is applicable to validating the subcomponents of the system. While traditional systems have test nodes to individually verify the operation of their subsystems, a classical test approach to SOP suffers from limited controllability and observability of its subsystems. Furthermore, the system specifications guaranteed by design depend on the validation of associated subsystem specifications, which may no longer be accessible in an SOP configuration. This proposition is especially important for embedded passives [6] constituting a part of the package. For example, a high-K resistor manufactured with a thin-film component embedded into the substrate may have a direct path to the substrate surface through microvias; however, once the associated application-specific integrated circuit (ASIC) is flip-chip bonded, there will be no directly accessible path for the ATE. Even if a separate design for testability (DfT) path is provided, the ATE will only be able to measure the combined effects of the resistor and the bonding parasitics of the ASIC. This example shows that HVM tests need to be partitioned across several steps of the integration process. Bare specifications of the ASIC and the substrate need to be tested before final assembly. Furthermore, the test set for the final system can be tweaked and reduced depending on those results. This multistep test strategy is discussed next.

The test flow for an SOP can be summarized in three steps: (i) known good die (KGD) test, (ii) known good embedded substrate (KGES) test, and (iii) known good embedded module (KGEM) test after assembly [7], as shown in Figure 5. The KGD test is performed separately on each bare die before assembly on the package. This test guarantees that bare dice or unpackaged ICs have the same quality and reliability as equivalent packaged devices [8]. More details about KGD tests, which are not addressed in this dissertation, can be found in [9–12]. Before attaching the bare die to the substrate, the substrate needs to be tested. Substrate testing includes

verifying the performance of interconnections [13] and the embedded passives [14, 15] fabricated in the layers of the substrate. Finally, the SOP is assembled and the final KGEM test is performed.

The call for a testable SOP results in a conflict of interest between the degree of integration afforded by the design process and the level of testability achievable by an external tester. A viable solution is to place the ATE functionalities in close proximity to the SOP module to be tested. This improves the test-access speed, minimizes test signal degradation, and increases controllability and observability of the signals internal to the DUT. One such candidate is the loadboard itself, where the test functions are migrated from the external tester to the additional circuitry built around the system under test. The additional circuitry retains the ability to apply high-speed stimuli to the system under test and capture the high-speed test response, which otherwise is degraded by the cable parasitics of the low-bandwidth external ATE. The resulting solution, called built-off test (BOT), presents a low-cost alternative to the prohibitive cost of a classical ATE. The other alternative, BIT, pushes the external tester functionality into the package and even into the bare dice wherever possible and is consequently a much more aggressive version of the BOT.

Note that without the BOT or BIT, very high-performance SOPs may not be economically testable. This is because the cost of external test equipment for test signal speeds in excess of 1GHz is very expensive. However, multi-gigahertz system designs are now becoming quite routine for high-bandwidth communications. The test economics is greatly improved by having high-speed test functions on the loadboard (BOT) or the SOP itself (BIT) augmented with low-bandwidth communication with a low-speed external tester. This allows high-speed systems to be tested with a low-cost external tester without loss of test quality.

Two BIT schemes are presented in this thesis. Both schemes depend on built-in

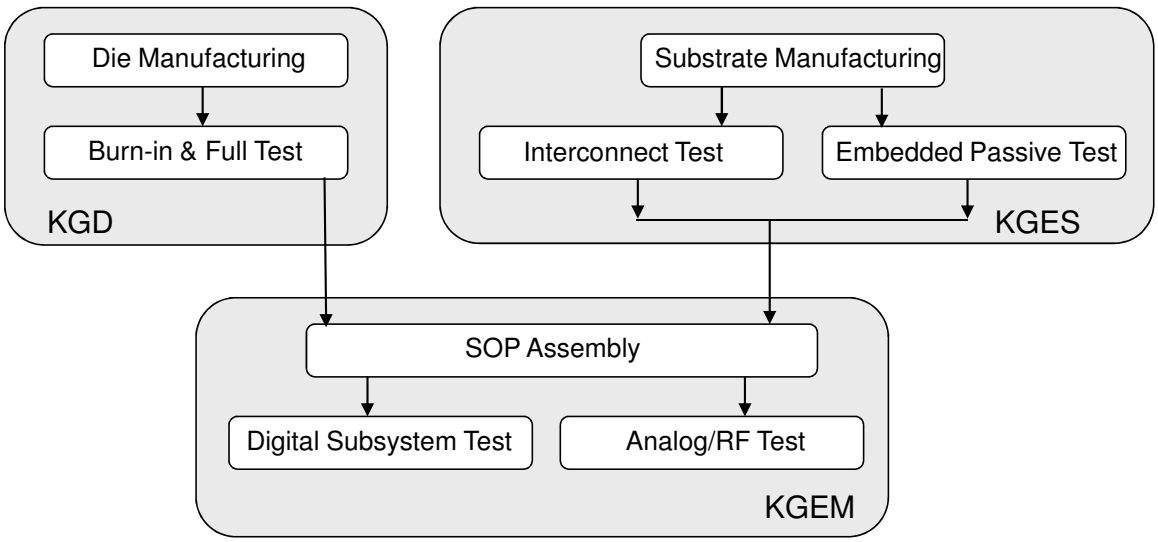


Figure 5: Test flow for SOP; known good die (KGD), known good embedded substrate (KGES), and known good embedded module (KGEM) tests.

feature extractors, which measure a complex function of the DUT response and output a simple signature. The first scheme, feature extraction with a noise reference, demonstrates the use of an alternate test for RF components embedded in a system with available DSP resources. The DUT is excited by an embedded oscillator, and a single-bit comparator captures the response. In chapter 4, it is demonstrated that this scheme can compensate for imperfect stimulus generation and performs well with low-speed and low-resolution comparators, which is ideal for low-cost and low-area overhead BIT implementations. In the second scheme, low-cost sensors are embedded into signal paths. The sensor characteristics are chosen in such a way that the low-frequency or DC signals at sensor outputs are tightly correlated with the target test specification values of the DUT. Hence, instead of testing the devices specifically for complex performance metrics, the outputs of the sensors are used to accurately estimate the target test specification values when the DUT is stimulated with a simple stimulus. The results in Chapter 4 show that the scheme not only predicts DUT performance accurately, but can also perform well under systemic process variations that impact the sensors as well, and under random variations in environment, such

as temperature.

1.3 Defect-Based Testing of RF Circuits

In the digital domain, effective test paradigms that rely on well-defined fault models, such as for stuck-at faults, provide reliable, abstract, and standard interfaces for powerful test algorithms to be developed and for their performances to be compared relative to each other. Similar fault models for analog testing have been proposed; however, the degree of abstraction in these models has been mostly device specific, and consequently, generic interfaces for test generation based on these models have not been developed. Although a generic fault-based test methodology is still the holy grail of analog testing, these devices are almost always tested in production for their performance specifications.

Existing fault models for analog/RF devices and associated coverage metrics are strongly dependent on the device and manufacturing technology. The common open, short, and bridge models only cover a small subset of possible defects. These defect-oriented catastrophic faults seriously impair the functionality of the circuit. The larger class, process-variation-oriented parametric faults, manifests itself as small deviations from the nominal circuit operating point. Most structural deviations in analog/RF circuits result in operating points well within acceptable limits. The differences between analog test paradigms come from the way they address deviations; either performance deviations can be tested against specification limits, or possible structural deviations can be eliminated against defect probabilities.

Defect-based testing, also called structural testing, works on the principle of negative elimination: each step of the test checks to see if a group of specific defects is present in the DUT; if the presence of a defect is detected, then the DUT is failed; otherwise it is shipped to the customer. The test generation scheme needs to have information about all kinds of possible defects, and will perform better if provided

with defect probabilities. This information is manifested as a defect list or as a defect dictionary. All catastrophic faults and large parametric faults are included in this dictionary. Process shifts are assumed to be detected by e-tests, and all affected wafers are discarded. Small parametric faults are assumed to be within acceptable performance limits by design or by proper redefinition of datasheet values. The main advantage is that most defects can be easily detected with very simple and cheap equipment. The ideal ATE for a defect-based test system is as simple as a voltmeter; as long as the voltmeter reading stays within predefined safe levels while predetermined input patterns are applied, the device is fault free. Although, applying a large set of these tests incurs long test time, multi-site testing can be applied to offset test time issues. In modern scaled CMOS processes, not all process variations can be effectively controlled anymore, and relaxing performance acceptance limits to accommodate such large process variability can negate the benefits of technology scaling in the first place. In addition, even in the presence of such large process variability, a significant percentage of dice still meet performance specifications. Throwing those dice out reduces yield. These problems constitute the main reason why modern analog test methodologies focus on testing against specification limits rather than eliminating defect probabilities.

The motivation behind chapter 5 is to selectively replace some features of alternate tests with those of defect-based test methods. By doing so, the ATE complexity and cost can be reduced by trading off alternate test complexity with parametric failure coverage with little or no impact on the coverage of physical defects. The resulting gamut of solutions has a common convergence point. The exact values of the specifications are not important although they may be obtained as a by-product of the alternate test procedure to varying degrees of accuracy. Redundant defects and minor process variations can be handled without taking a hit on yield. It is also not necessary to have a detailed fault list for defect-based testing as long as the

boundaries of specification violation under dominant fault mechanisms are known. Three of these methods are studied: (i) signature filter, (ii) alternate support vector machines (SVM), and (iii) dictionary-independent SVM.

The simplest extension of an alternate test to address faults is to divide the mapping process into two steps. The first step is to apply a coarse filter, a defect filter, to analyze the DUT response marking out catastrophic faults, followed by a fine mapping from the DUT response to the DUT specification values for parametric faults. The second step is to abandon specification prediction and simply stick with a go/no-go decision. This can be achieved by using classifiers instead of regression methods. It is demonstrated that with a test generation algorithm based on SVM, (i) the go/no-go approach has a significant effect on ATE complexity compared to the regression-based mapping approach, and (ii) e-test data can be used to improve the performance of this approach in terms of fault detectability. An extension of this method is to improve the go/no-go approach such that the final fault-based alternate test is defect-dictionary independent. Only Monte Carlo (MC) simulations of process variables are used for training, with no hard defects in the training set. The stimulus optimization is driven by an artificial go/no-go boundary for the SVM, which is gradually progressed from the zero diameter centered at the optimal design to the diameter enclosing all Monte-Carlo simulations. The stimulus search is performed in the direction that best represents the progress in the corresponding measurement space. Each of these three methods represents a different trade-off point in test cost, flexibility, and implementation complexity, and together they provide a solution matrix for products with different market profiles.

Chapter II

SOP TEST METHODS AND ALTERNATE TEST METHODOLOGY

RF circuits can be stand-alone, such as amplifiers or mixers that can be bought off-the-shelf, or they can be manufactured as a part of a bigger system, such as SOPs. Test constraints are significantly different for these two cases; consequently, an RF test solution that works well for a stand-alone component may be sub-optimal for an RF circuit buried in an SOP. Before going into details of RF test methods, fundamentals of fault models, test quality metrics, and test techniques complementary to RF components in a system are studied in this chapter. Following these concepts, research literature on RF testing is summarized and a survey of alternate test methodology is presented.

2.1 Fault Models and Test Quality

Faults in analog, mixed-signal, and RF circuits are broadly classified into two categories [2]: catastrophic, in which the component fails to operate correctly as a result of internal manufacturing defects such as shorts and opens; and parametric, in which one or more specifications of the device deviate from the respective design values as a result of variations in the manufacturing process. Defect-oriented tests (DOTs) [3] are based on finding a suitable test signal to detect the presence of catastrophic faults using different automated fault simulation and test generation techniques [16–20]. The specification-oriented tests (SPOTs) [21–25] are concerned with a direct or indirect measurement of the specification on the device datasheet. Under these two categories, test quality metrics are defined so that the effectiveness of a test methodology can be

evaluated and these metrics can be directly compared for various test methodologies.

While fault coverage [26] is an accepted test quality metric used for testing digital circuits, its extent and meaning in the analog domain are not completely clear in the literature. Often, the analogy between stuck-at faults in digital circuits and opens/shorts in the analog domain is carried too far into the fault coverage of an analog test and is defined as the percentage of potential shorts and opens the analog test can detect. However, catastrophic failures that result in significant performance loss of the DUT can be detected by simple tests, especially in RF components, where there is usually one or very few independent signal paths, and the signal power at the output is significantly smaller when there is a catastrophic failure. In reality, the effectiveness of an analog test methodology is largely dictated by its ability to detect parametric failures of the DUT where performance deviates by a small amount from the nominal. These parametric failures are more likely to occur than catastrophic ones but are harder to detect than the latter. Furthermore, the meaning of a parametric fault is not clearly defined since any excess variation in a component's value, although considered a fault, may have little impact on device specifications. If the test methodologies geared toward increased fault coverage, especially defect-oriented ones, base their evaluation on parametric faults of individual components, they will eventually end up compromising yield coverage, the probability that a fault-free device passes the final test [27].

In BIT applications, an important test quality metric is the area overhead, the percentage of extra area introduced by testing related electronics. It is a major concern for practical implementations since this extra area does not add any value after the device passes the production test barrier. This argument, however, is not valid for some built-in self-test (BIST) solutions in which the test can be applied throughout the life span of the product. Often the area overhead is overemphasized when compared to the yield coverage in the qualification of BIT methodologies. However,

only the joint figure of these two metrics can define the effective wafer area dedicated to the product.

Any system-level test methodology needs to ensure that all the specifications in the device datasheet are verified in production before the device is shipped to a customer. One possible approach to achieve this objective is testing every individual submodule of the device followed by testing of the proper connectivity of the submodules inside the SOP. In effect, this approach breaks down the system testing problem into many smaller module-level testing problems. Although this approach requires physical test access to the individual internal submodules for module-level testing, it is often more effective than end-to-end system-level testing in terms of both the production test feasibility and the test cost. One such example is wireless transceiver applications, where the testing of RF signal blocks, intermediate frequency (IF) signal blocks, and codec blocks can be performed independently, and the connectivity of the concerned modules can be verified subsequently to qualify the devices as "good" or "bad" in production tests.

In such bottom-up test procedures, algorithms for relating the individual submodule test responses to the system-level test specifications of the SOP must be devised to aid in the pass/fail decision making process. The key is that any circuit, submodule, or system-level failure that causes any of the system-level test specifications of the SOP to be violated is defined to be a fault. A valid test methodology must be designed such that it can detect even the smallest of manufacturing defects that can result in such a fault. If no suitable algorithms can be found for determining the system-level test specification values from the SOP submodule test responses, then the only recourse is to directly measure the relevant test specifications at the system level. In general, this is more expensive than running submodule-level tests. A typical example is testing of the input referred third-order intercept point (IIP3) specification of circuits exhibiting nonlinear behavior. Measuring IIP3 end-to-end

requires high performance, that is expensive, measurement instruments. However, if the specification can be inferred from the results of submodule test, then the same test can be performed using a simpler setup [28].

2.2 Substrate Interconnect Tests

Package interconnections consist of single or multiple layers of metalization that connect active circuitry and embedded passives to form a function. If economically viable, the various interconnection layers can be optically inspected during the processing of the layers for the presence or absence of conductive material along the interconnection length. This allows for the immediate detection and repair of process-related defects during fabrication. In very high-density interconnect (HDI) solutions, typical in SOP, the optical inspection is not feasible, so the methodology relies on good process control. Even when every layer is optically inspected, temperature and process stressing of subsequent layers lead to defects on interconnections that need to be diagnosed prior to die attachment. Hence, a test, which cannot be done through optical inspection, is required after all the layers are fabricated.

In the older paradigm, with laminated substrates and plated-through holes, the substrate test can be easily performed through the top and bottom surfaces of the substrate by electrically probing the interconnections. In SOP solutions, HDI with microvias is the norm, and this method mostly creates blind and buried interconnect layers that reach neither of the substrate surfaces. Surface-to-surface, as well as ground and power interconnects can still be tested by probing both exposed ends, as shown in Figure 6a. Some interconnects connect a surface contact to an embedded thin-film passive, which is in turn connected to a power or ground net, as shown in Figure 6b. Testing of these interconnects can be performed in a single step with the embedded passive they are connected to. Other interconnects, as shown in Figure 6c, connect an embedded module to another embedded structure or to power/ground.

These interconnects can be tested in a single step with the embedded modules.

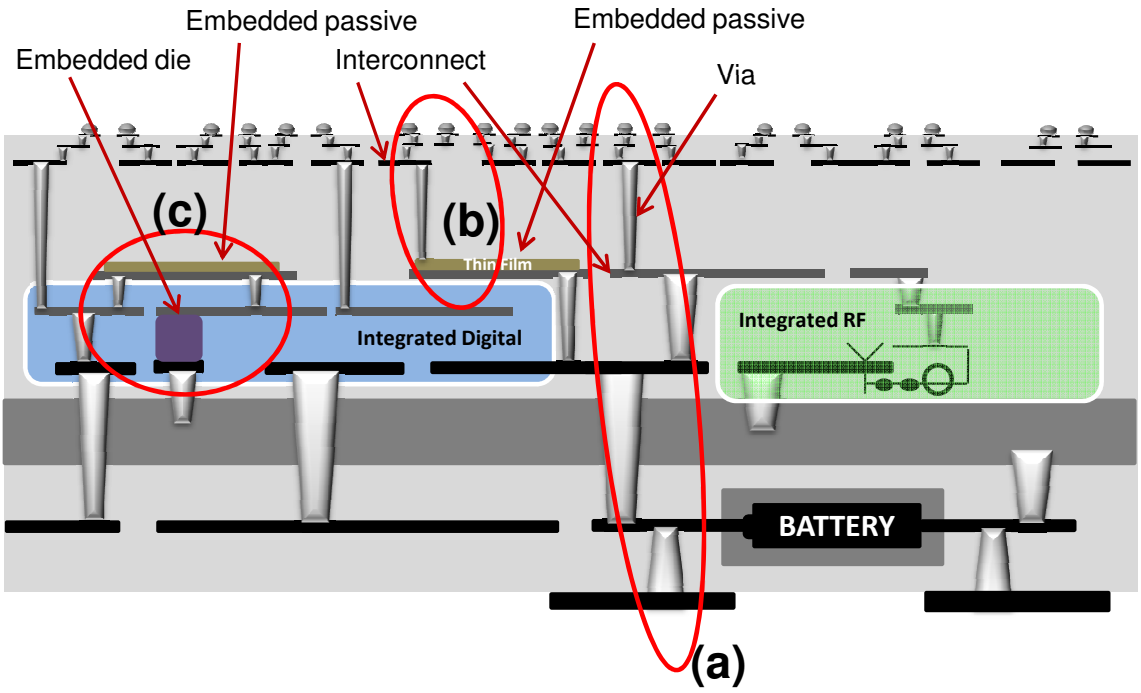


Figure 6: SOP interconnect test (a) interconnects exposed on both surfaces, (b) an embedded passive with one end connected to power or ground, and the other end exposed, (c) an embedded passive connected to an embedded die.

2.3 Testing Embedded Passives

With advancements in SOP manufacturing technology, embedded passive components are playing a key role in the development of these systems. The integration of passive components into a system offers inherent benefits, such as a reduction in the size of the system, a reduction of parasitic effects, lower assembly cost, and superior electrical performance. However, testing these embedded integrated passives is more difficult because of the inaccessibility of internal circuit nodes, the possibility of many different failure modes resultign from the analog nature of these components, and the high frequencies of circuit operation especially for RF passives. Three important parameters critical in the design of embedded passive components are [29]: (i)variation of

resistance and reactance with frequency, (ii)variation of the quality factor of the components with frequency, and (iii)the resonance behavior of the components. The main research in testing of embedded passives originates from advancements in multi-chip modules (MCMs) [30–34]. State-of-the-art methods are high-frequency at-speed tests, which require measurement of network parameters as a function of frequency [6].

2.4 KGEM Test of Digital Subsystems

The testing of digital subsystems is mostly standard, involving a boundary scan. A boundary scan is a test interface standard that defines the way digital data can be serially scanned into and out of any package-level system [35]. This structure enables controllability and observability in large complex digital systems. The basic structure of the boundary scan is illustrated in Figure 7 . Here it is assumed that each I/O cell has additional test logic beyond that needed for normal system functionality. Each cell has a serial input and a serial output, and by connecting the output of one cell to the input of another, the scan chain is formed. During test mode, this additional logic allows data to be serially shifted from one cell to another. At one end of the chain is an unconnected serial input called test data in (TDI). At the other end of the chain is the test data out TDO). In order to clock test data through the chain, a test clock TCK), which is separate from any other system clocks that might be needed for normal functions, is used. A test mode select TMS) signal is used to control the switching of the scan logic between scan mode and normal-system mode. There is the optional test reset TRST*), which is used to reset the test control logic and to deactivate the boundary scan register.

A scalable modification of the boundary scan has been successfully applied to testing at the system level. The classical paper from Zorian [36] lays out the structured testability approach, as shown in Figure 8, which is well accepted in industry. Although a mixed-signal boundary scan standard also exists, analog test stimulus

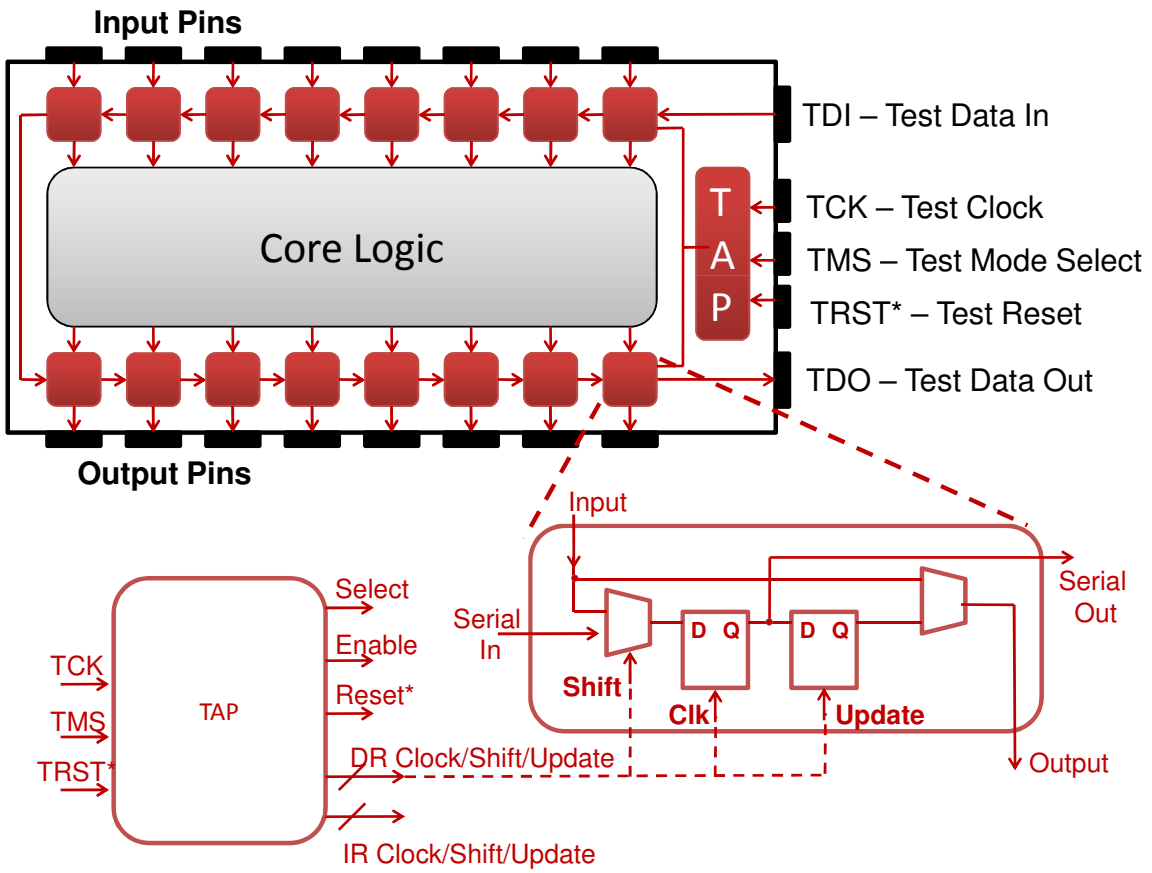


Figure 7: Boundary scan structure.

generation and response acquisition remain as open problems [37]. The main challenge for boundary scan testing of SOPs is the reliability of the solution. As the number of dice increase, the simple scan chain gets longer. Motivating the test hardware design is the need to test the stack electronics reliably in the presence of faults in the boundary-scan circuitry, such as an open TDI line. The work in [38] evaluates possible scan test methodologies in terms of test reliability, scalability, and testing time overhead introduced by the system-level approach. The best trade off is a partitioned approach with separate multi-chip test controllers integrated into the SOP. Redundant test lines enable testing in the event that some lines are faulty.

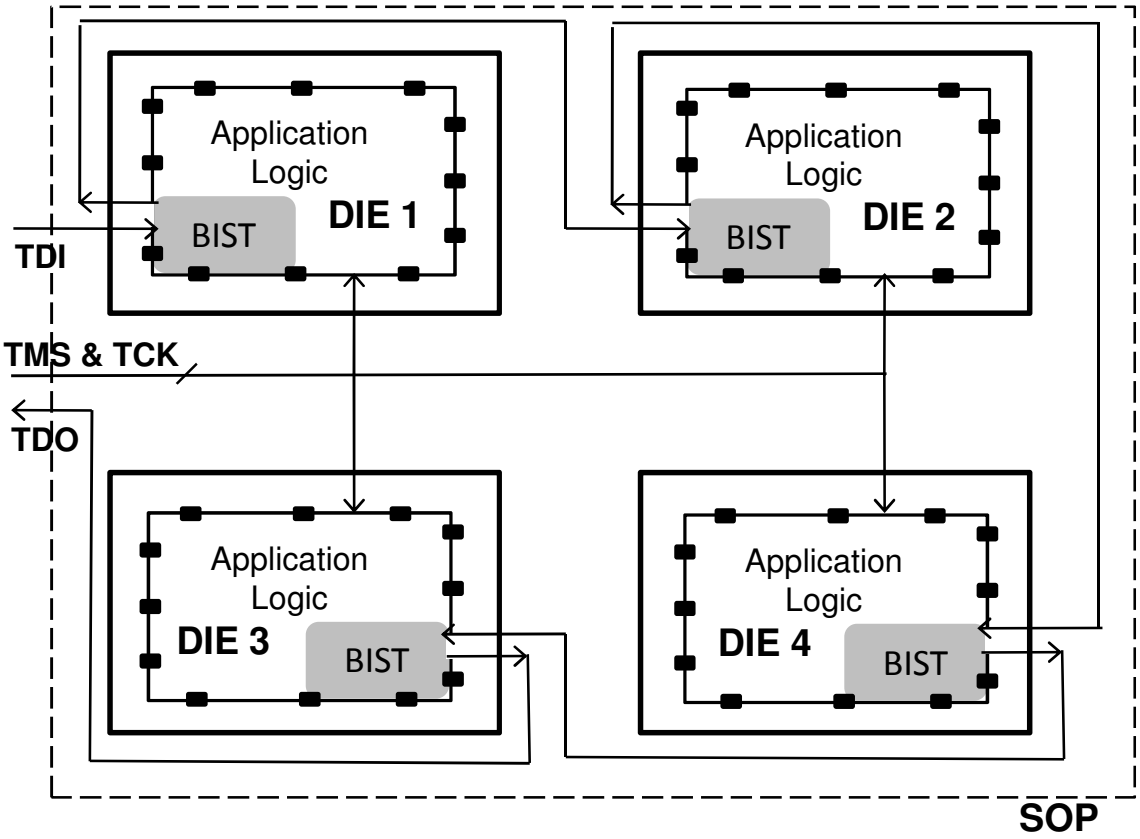


Figure 8: System-level boundary scan.

Another challenge with digital subsystems is the multi-gigahertz IO. The characteristic size of high-speed digital ATE scales inversely with the maximum IO frequency. At frequencies above 10GHz, the scaling trend suggests that critical test electronics must be smaller than the chip, implying full BIST. A detailed survey of high-speed digital test literature is available by Akbay et al. [39]. State-of-the-art contributions focus on pushing driver and sampling capability closer to the DUT [40], and enabling wafer-level test solutions in the form of test support processors [41].

2.5 Literature on RF Test Generation

In the digital domain, there is a large amount of research published for automatic test pattern generation (ATPG) and BIST. In this section, only the studies related to analog and RF components are reviewed. Previously developed approaches can be

studied in seven main areas.

2.5.1 Test Selection and Ordering

These techniques are usually limited to external ATE rather than BIT. In this approach, the coverage metrics for each of the specification tests are evaluated individually and tests are ordered in such a way that the average testing time is minimized. The optimum order is likely to start with the test that can cover the largest percentage of possible faults and continues until all tests are applied or one test fails the component. The optimum order may be different when there is a large number of specifications to be tested and tests using the same setup can be grouped together to eliminate extra switching time. Test time reduction is obtained only when a fault is detected, and the rest of the tests are terminated. However, for a mature HVM environment, only a small percentage of DUTs are faulty to begin with; hence, the test time savings is usually marginal and may not be substantial enough to justify the extra engineering time spent on test ordering. On the other hand, some specification tests may be eliminated altogether if the remaining ones can supply the original test coverage. Similarly, it may be possible to reduce the complexity, and consequently the cost, of loadboards. It may even be possible to cover this reduced number of tests with a cheaper ATE. Note that the cost benefits quickly diminish if a binning scheme is required rather than a simple pass/fail decision. A brief literature survey of these methods is provided in Chapter 3.

2.5.2 Multi-Tier Testing

Test ordering provides only marginal test time reduction because the tests are ordered from a negative point of view, i.e, their likelihood to fail a component. Multi-tier testing is based on a positive point of view. The original set of tests is replaced by a set of cheaper tests that can make a pass/unknown decision rather than a pass/fail decision. Only if this first tier of tests returns an unknown is a second

tier of more expensive tests performed. Higher tiers can be added as long as the extra cost of logistics is justified by the provided test cost reduction; otherwise, the components labeled unknown by the highest tier are discarded. The average test cost is a function of the test cost for each tier and the probability of each tier catching a fault-free device. The main downside of multi-tier testing is that the critical first tier of tests is determined in an ad-hoc manner, and the possibility of finding such tests is highly component specific. Furthermore, unless all tiers are performed on the same loadboard and ATE, the extra handling time required from one tier to the higher is likely to be prohibitive. A recent and high profile example is the missing tone power ratio (MTPR) test. A literature survey of these methods is provided in Chapter 3.

2.5.3 Defect-Oriented Testing (DOT)

DOT is based on the assumption that most or all fault mechanisms manifest themselves in more fundamental observables than the specifications. These observation mechanisms can be discovered by extensive analyses of the DUT through simulations. Empirical data is collected on a large set of DUT samples to build the required fault models. Some DOT solutions serve as the first tier in a multi-tier approach. A typical example is the *Iddq* test, for which the leakage current of the DUT is measured and classified. Although *Iddq* used to be a very common screening tool for HVM tests of digital circuitry such as microprocessors, its efficiency has been diminished by the introduction of submicron processes. Modern extensions of *Iddq* such as *Iddt*, *delta-Iddq*, and *min-Vdd* are subjects of active research. Similar tests are employed for power-hungry RF components such as power amplifiers (PAs). The main disadvantage is that the success of DOT heavily depends on fault models and fault lists, which can be assessed accurately only in the later stages of the HVM life cycle. A brief literature study of these methods is provided in Chapter 5.

2.5.4 Design-for-Testability (DfT)

In DfT, additional hardware or software features are implemented at the design phase such that the debugging and/or HVM testing is simplified. A common DfT feature is the extra input and output ports that improve the controllability and observability of the intermediate stages. Note that BIT solutions are extreme DfT features. Although DfT is a common tool in the digital practice, it is harder to justify extra ports for RF components for several reasons, including sensitivity to noise, power loss and area overhead of the required signal couplers, and more complicated matching constraints. A brief literature survey is presented in Chapter 4.

2.5.5 Built-Off Tests (BOT)

BOT provides a compromise between the ATE solutions and BIT by moving some of the ATE functionality onto the loadboard. It provides a work-around for devices that presents a mismatch between the bandwidths available outside and inside the chip, yet the design can no longer be modified with BIT structures. Although complicated loadboards are still much cheaper to manufacture than high-performance ATEs, when repeatability and reliability come into play, the total cost may be comparable. Furthermore, since ATEs are generic tools, the related engineering investment is a one-time cost component, whereas loadboards need to be designed almost from scratch for different DUTs and ATEs. A brief literature survey is provided in Chapter 4.2.3.

2.5.6 Direct Measurement of Specifications Using Dedicated BIT Circuitry

In this approach, the external ATE functionality is designed inside the DUT for applying appropriate test stimuli and measuring the test response corresponding to the SUT. The scheme may still require a low-bandwidth channel between the BIT hardware and an external ATE to complete the test. Although the direct measurement

procedures are conceptually simple, their implementation has two major drawbacks: (i) multiple specification measurements require different kinds of resources, which add up to significant area overhead; and (ii) a longer total test time is required because multiple specification measurements cannot be performed simultaneously. A literature survey is presented in Chapter 4.

2.5.7 Specification-Based Alternate Tests

In this approach, the datasheet specifications of a DUT are statistically calculated by analyzing DUT response to a specific test input. This stimulus is carefully crafted to yield a significant correlation between the response and the specification variations. The DUT response can be considered a signature for the effects of process variations on that particular DUT instance. Supervised learning techniques are used for constructing the mapping functions that can output the best estimates for all specification values given a DUT response signature. Alternate tests can significantly cut down testing time because all specifications are calculated from a single test setup and a single test stimulus. For the same reason, the loadboard complexity and cost can be reduced. Alternate tests promote low-cost ATE solutions where the ATE functionality is tailored to accommodate the simple alternate test resources, or an old ATE with limited resources can be used for testing modern RF components. Alternatively, the resources available on a state-of-the-art ATE can be partitioned to simultaneously test multiple components. This is possible because alternate tests require only a single configuration on the loadboard, making more estate available for multiple test sites. The disadvantage of alternate tests is that they require a large calibration set of devices for the mapping functions to cover all possible process variations and catastrophic faults. The fundamentals of specification-based alternate tests are discussed next in this chapter. All presented research in this dissertation makes use of alternate test techniques; hence, brief literature surveys are presented in each

chapter.

2.6 Alternate Test Primer

In recent years, a growing number of publications on alternate tests have proposed new solutions for the analog/RF test bottleneck. In the alternate test approach, a suite of sequential specification tests is replaced with a single test consisting of a carefully crafted test stimulus applied to the DUT. The response of the DUT to the applied alternate test stimulus can be mapped to all SUTs concurrently [42], thereby allowing significant test time savings. Simultaneously, the cost of the external tester can also be reduced since a simpler test setup can be used to measure all different test specification values of the DUT.

The first implementations of alternate tests for low-frequency analog circuits made use of a time-domain oversampling procedure to generate the response samples. However, oversampling poses a great challenge in the RF domain. Subsequently, another variant based on the use of upconversion/downconversion mixers to change the frequency range of the stimulus and response [43] was proposed. In order to address the sampling problem, [44] proposes the use of sensors that convert the obtained RF response to DC-level signatures. A combination of BIT and BOT alternate methodologies is explored in [39], especially for SOPs and systems-on-chip (SoCs), for which observability and controllability are the major problems. The use of generic alternate test modules enables limited on-chip hardware to perform the test functionality of many specialized modules. This idea is further explored in [45, 46] for BIST implementations and in [47] for on-chip self-calibration.

In specification-based alternate tests, the datasheet specifications of a DUT are predicted by analyzing its response to a specific input pattern. This stimulus is carefully crafted to yield a significant correlation between the response and the specification variations. The DUT response can be considered a signature of the effects of

process variations on that particular DUT instance. This procedure is shown in Figure 9. The many different stimuli in the standard specification tests are replaced with a single alternate test stimulus; however, specifications can no longer be extracted by simple mathematics. Nonlinear statistical multivariate regression analysis [48] allows one to construct mapping functions such that for a given set of measurements, these functions generate predictions for the values of each specification [42].

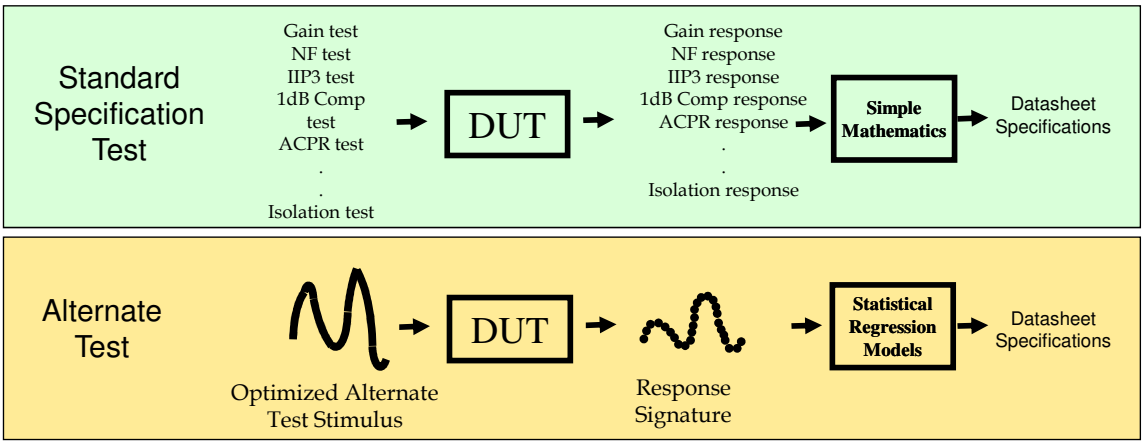


Figure 9: Standard specification test scheme vs alternate test scheme.

The key to an accurate specification prediction can be summarized in three principles: (i) the alternate test stimulus is selected to maximize the correlation between the response and specifications under test, and hence requires elaborate ATPG algorithms; (ii) the response signature provides a robust basis for the mapping functions to convert the single signature into many specification values; (iii) the mapping functions are generated by a supervised learning process on a sample set of training devices for which the specifications under test are measured with conventional test methods.

The alternate test generation flow can be studied in four steps: (i) craft an alternate test stimulus; (ii) calibrate the mapping functions using hardware measurements on a sample set of devices; (iii) apply the alternate test in HVM; (iv) update the sample set and go back to step (ii) when process screening indicates a significant shift from the characteristics of the sample set of devices used in initial calibration.

Figure 10 shows the details of a typical alternate test flow. On the left, standard specification tests are performed on the sample set and actual specifications are recorded, the sample set is divided into a training set and a validation set. On the right, the alternate test scheme makes use of the training set to find an optimal test stimulus, and the corresponding signatures are collected. The actual specifications of the training set together with their alternate signatures are fed into a knowledge discovery process, yielding the mapping functions. These mapping functions are used to predict the specifications of the validation set. Finally the difference between the actual and predicted specifications gives an accuracy estimate for the proposed alternate test.

The specification values for each device are obtained by mapping the signatures into the specification space. These mapping functions are constructed by a supervised learning process called multivariate adaptive regression splines (MARS) [48]. The final functions can be visualized as a weighted sum of basis functions made of splines, which span values for each of the independent variables. The mapping function f for a specification y with M elements in the signature can be represented as

$$y = f(x) = \beta_0 + \sum_{m=1}^{M} \beta_m B_{km}(x_{v(k,m)}), \tag{1}$$

where the summation is over the M independent variables and β_0 and β_m are parameters of the function. The basis function B is defined as

$$B_{km}(x_{v(k,m)}) = \prod_{k=1}^{K_m} H[s_{km}(x_{v(k,m)} - t_{km})], \tag{2}$$

where $x_{v(k,m)}$ is the k^{th} independent variable of the m^{th} product, K_m is the number of splits that gave rise to B_m, s_{km} can be ± 1 depending on the right or left sense of the step function H, and t_{km} are the knot locations for these step functions. MARS uses an initial recursive partitioning during training to gradually add these basis functions using forward stepwise placement; then, a backward procedure is applied

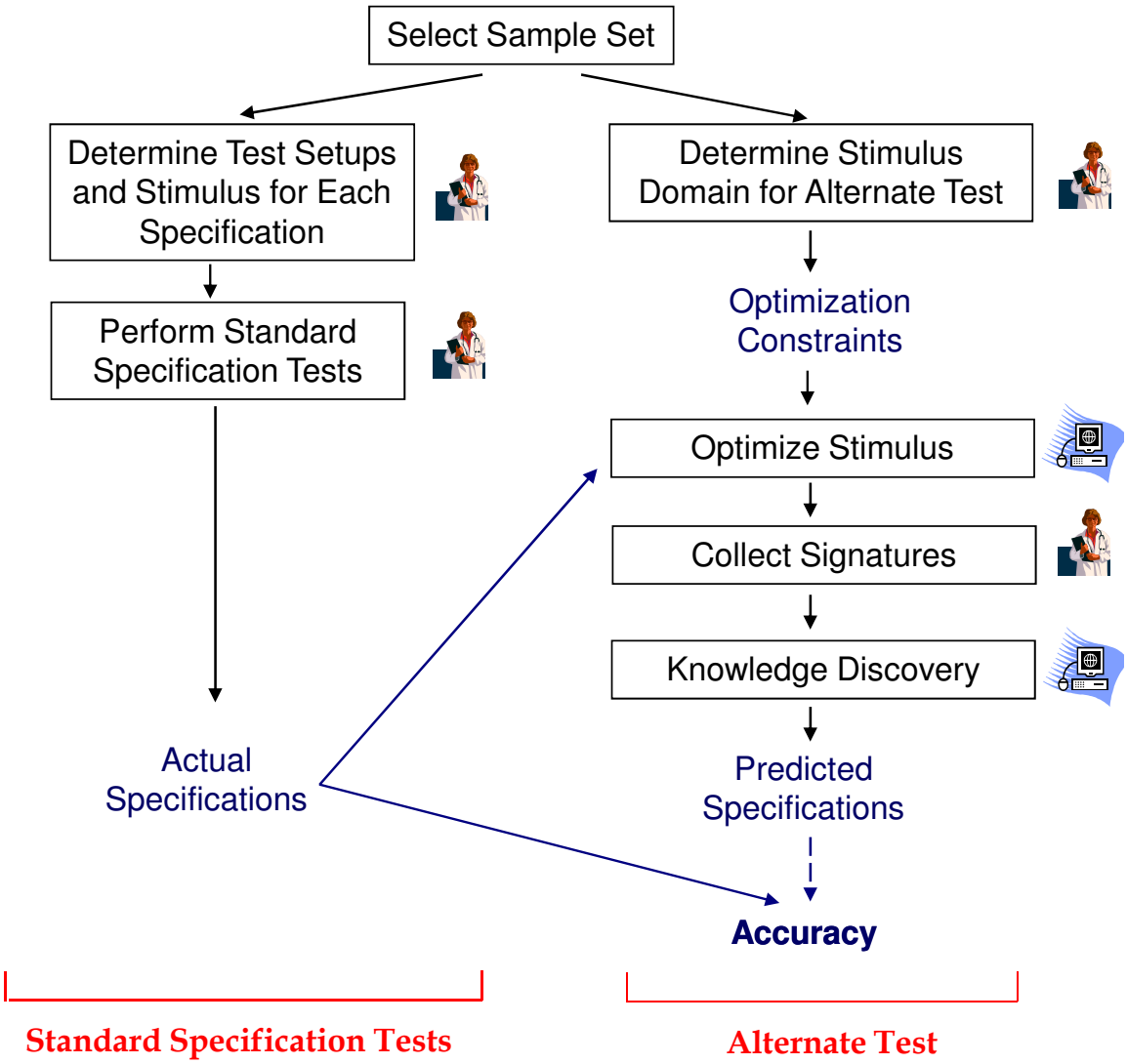


Figure 10: Evaluation scheme for alternate tests.

and the basis functions associated with the smallest increase in the least squares fit are removed.

Chapter III

CONSTRAINT-DRIVEN RF TEST STIMULUS GENERATION

In this chapter, three test stimulus generation algorithms are studied. Each one corresponds to significantly different constraints, ranging from trivial to challenging, while moving from a domain-specific algorithm to a generalized methodology. The first example, multisine tests, works in a domain that provides complete access to circuit netlist, and data about process variations in manufacturing. The associated search algorithm, MUSTI, is a memoryless local optimizer with a greedy decision mechanism. This algorithm demonstrates a highly-tuned solution to the domain-specifics of the problem and provides a baseline for the generalized methodology by highlighting short-comings of a specific trivial solution. The second example, current signature test, implements a BIT-friendly solution that is not invasive to the signal path. This example demonstrates the use of different signatures with a simple alternate test stimulus. Tradeoffs are studied by comparing a voltage-based signature with a current-based one. The third example, test generation with behavioral models, demonstrates a generalized test generation methodology. It works in a domain that does not provide access to netlist or process variation details. The high-level models of the test signal path are used to define constraints in the search space of the optimized test stimulus. A fine balance is pursued, where these models are generated in enough detail such that they inherently define limitations of the low-cost ATE and the I/O restrictions of the DUT, yet they are simple enough that the nonlinear optimization problem can be solved empirically in a reasonable amount of time. Various search algorithms and signature mapping techniques are compared.

3.1 Multisine Test Stimulus Generation with Greedy Search

The first approach evaluates a greedy search algorithm based on merge/split operations in the frequency domain. The test generation algorithm optimizes a single multitone sinusoidal input waveform in such a way as to detect all manufacturing defects and process variations that also affect the test specifications of the DUT. The cost function is defined in terms of several design variables, including maximum input frequency, test time, and test coverage. The multitone transient input is a superposition of a number of sinusoidal waveforms. The algorithm uses a single transient input to test for multiple specifications at a time. This is a very important point for reducing the test time; only one waveform is required for testing all the specifications. Furthermore, the waveform generator is relatively simple in complexity compared to an arbitrary signal generator. There are two key aspects of the algorithm: (i) the use of an optimization method to limit the number of tones and the highest frequency component of the multitone input, and (ii) the use of a controlled experimentation method to explore the parameter space in an efficient manner.

The implementation example is a 900MHz low-noise amplifier (LNA) [49]. The test results on this LNA suggest that with the presented test generation algorithm, it is possible to have high fault coverage at test speeds lower than the nominal operating frequency, and within a reasonable test time. Furthermore, specifications that are typically measured by applying a superposition of two sinusoidal tones, such as IIP3, are shown to be effectively measured by a single-tone input waveform. The actual versus predicted specifications for 1dB-compression point (1dBComp), IIP3, and the noise figure (NF) are shown for the LNA example. The findings are published in [50].

3.1.1 Background

Fault-based testing is an interesting alternative to functional testing in the RF domain. In this methodology, the target of test generation is to maximize the difference between the fault-free and faulty circuits. Depending on the application, researchers have employed various types of test inputs. To name a few, in static DC testing [18, 51], a DC voltage or current is applied; whereas frequency domain testing [52, 53] uses sinusoidal signals to study the steady-state response of the circuit; transient testing [54, 55] applies piecewise linear or multitone voltage waveforms and samples the transient response.

In [18, 51–55], the testing method is evaluated in terms of pass-fail analysis with regard to "hard" catastrophic failures. In [42, 56], a new methodology is proposed to predict "soft" performance parameters directly from transient testing. The approach in this section, focuses on parametric testing of RF amplifiers and evaluates its performance with "hard" boundaries that represent margin of test acceptance.

3.1.2 Proposed Approach

The multisine waveforms are optimized in terms of a cost function to limit the number of tones, the maximum test frequency, and the total test time by using constrained optimization. As a result, the tests are done at the lowest possible frequency satisfying maximum test time limitations.

A majority of current RF architectures allow access to individual components of the design through external points. Figure 11 shows the corresponding test setup for a RF device. The ATE generates a multisine transient waveform, which is the input for the RF DUT. The ATE samples the output of the DUT, and the specification mapping module (SMM) uses these voltage samples to generate predictions for multiple specifications of interest at a time. The comparator checks the predictions against specification pass/fail tolerances, and generates an output to identify whether

the DUT fails for any one of these specification.

The proposed algorithm, multisine test input optimizer (MUSTI), takes (a) the netlist of the DUT, (b) device models, (c) process variables and distributions of process variables, (d) nominal values of specifications, (e) specification pass-fail tolerances, (f) nominal operating frequency of the DUT, (g) quality measure for prediction accuracy, and (h) the accuracy of ATE, as inputs; and generates (a) the information necessary to build a multitone input transient waveform, (b) the sampling interval, sampling start time, and number of samples, (c) a specification prediction model for each specification of interest, as the outputs.

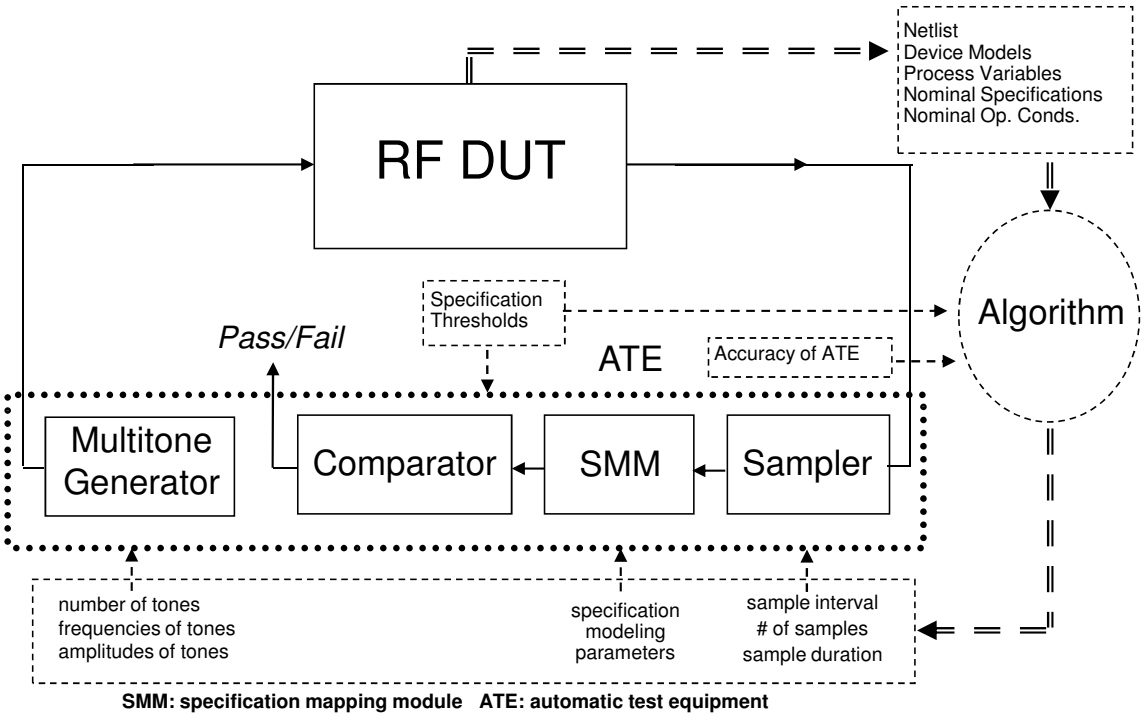


Figure 11: The proposed test setup for a component within an architecture that allows access to individual components.

Figure 12 shows the control and data flow for the algorithm. The upper part is computed only once for a given RF device. Two sets of device instances are generated for training and validation purposes using the circuit netlist, device models, and process variable distributions. SpectreRF simulator is used to process all of instances at

the nominal operating frequency and at the nominal temperature of operation. These simulations are designed to measure actual specifications of interest for each circuit instance. The lower part of the figure describes the iterative optimization loop. First, a controlled experimentation methodology proposes a number of possible simulation and modeling sets. Each set is a collection of the following entities: the number of sinusoidal tones, frequency and amplitude of each tone, sampling interval, sampling start time, number of voltage samples, and number of basis functions for a specification model. A number of these sets are selected considering the limits determined by the constraint optimizer. The first six entities of every set describe a transient analysis, simulations that correspond to the selected sets are run in parallel. The result is a set of sampled voltage values for each simulation. The voltage samples of the training set are used with the number of basis functions, the accuracy of the ATE, and the actual value of specifications to generate a MARS model for each specification. Then, these models are used with the voltage samples of the validation set to generate predictions of validation set specifications. The predictions are compared with actual values, and pass/fail thresholds are applied to calculate the accuracy of predictions, which are stored in a database. The constraint optimizer uses this database to select new bounds on number of tones, maximum frequency value, and maximum testing time. If the optimizer satisfies all the objectives, optimization is over, if not, this loop is executed again with the new bounds.

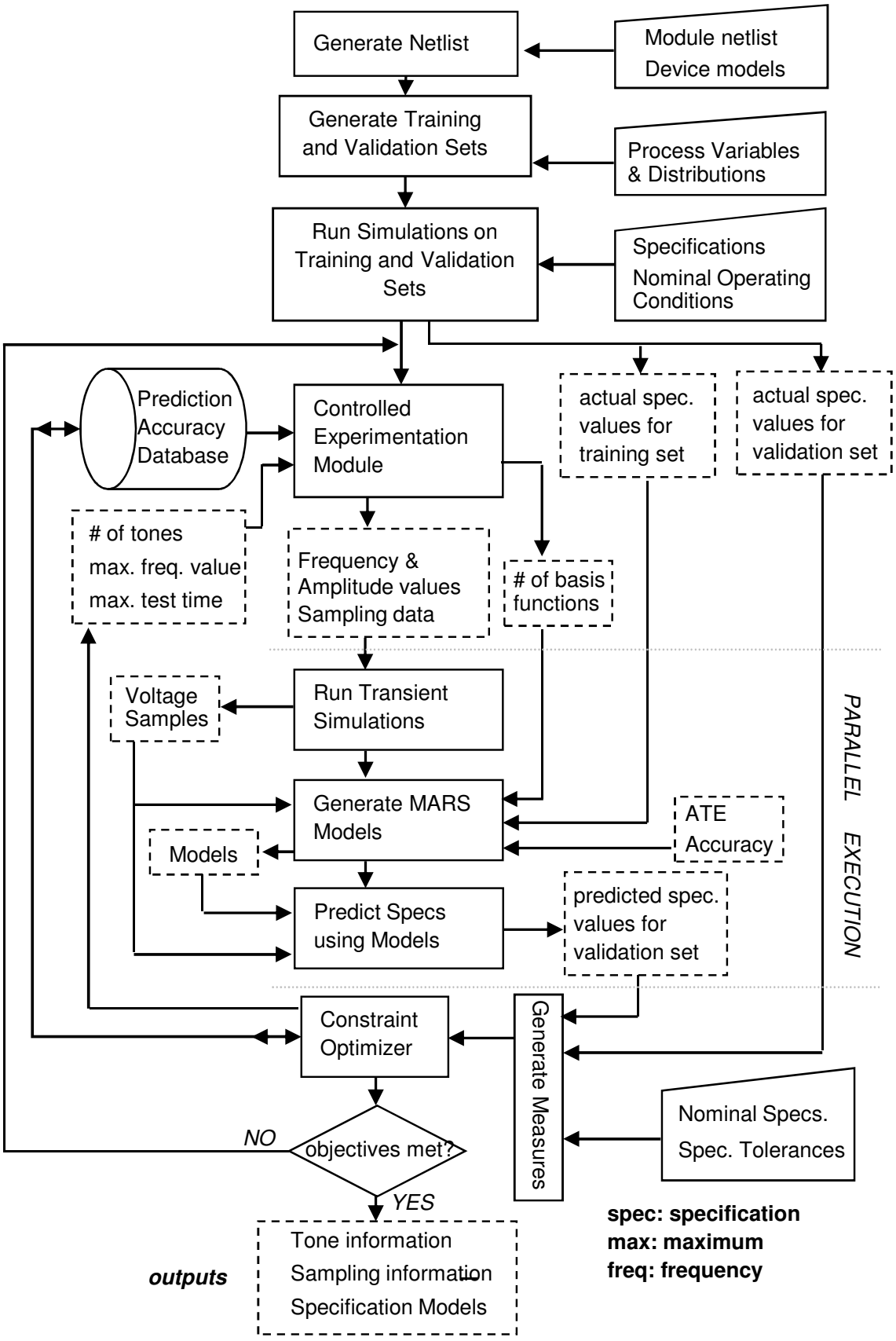


Figure 12: Control and data flow in MUSTI algorithm.

3.1.3 Optimization Scheme

Optimization may be applied to various stages of the proposed testing methodology, including input stimulus generation, simulation parameters, sampling parameters, and mapping functions. Practical concerns limit the use of optimization in some of these parameters. The parameters like simulation accuracy, sampler capacity, the distribution of process variables are usually defined by a standard ATE implementation and supporting framework. When the optimization model is based on an explicit relation, as shown in Figure 13, the surface fitting module proposes a number of experimental parameters for the determination of the unknown parameters of the explicit function. These parameters are converted into initial experiments, and the corresponding multisine inputs go through the simulation process. Then, the surface fitting module solves for the unknown parameters using the prediction accuracies of these simulations. The final form of the function is evaluated by the optimizer, and the corresponding multisine input goes through the simulation cycle once more to produce the final models for each datasheet specification. These models and the final multisine input are the outputs of the procedure.

The relation that is made explicit by the surface fitting module may be solved by analytical optimization methods. When the relation is expressed in terms of a multidimensional function, the optimization procedure solves for the zero-crossing of the gradient of the function, and searches for a stationary point. The nature of this stationary point is examined by evaluating the Hessian of the function. When there are further constraints in the independent variables of the function, these constraints are associated with Lagrange multipliers and the optimization is performed relative to both the independent variables and Lagrange multipliers.

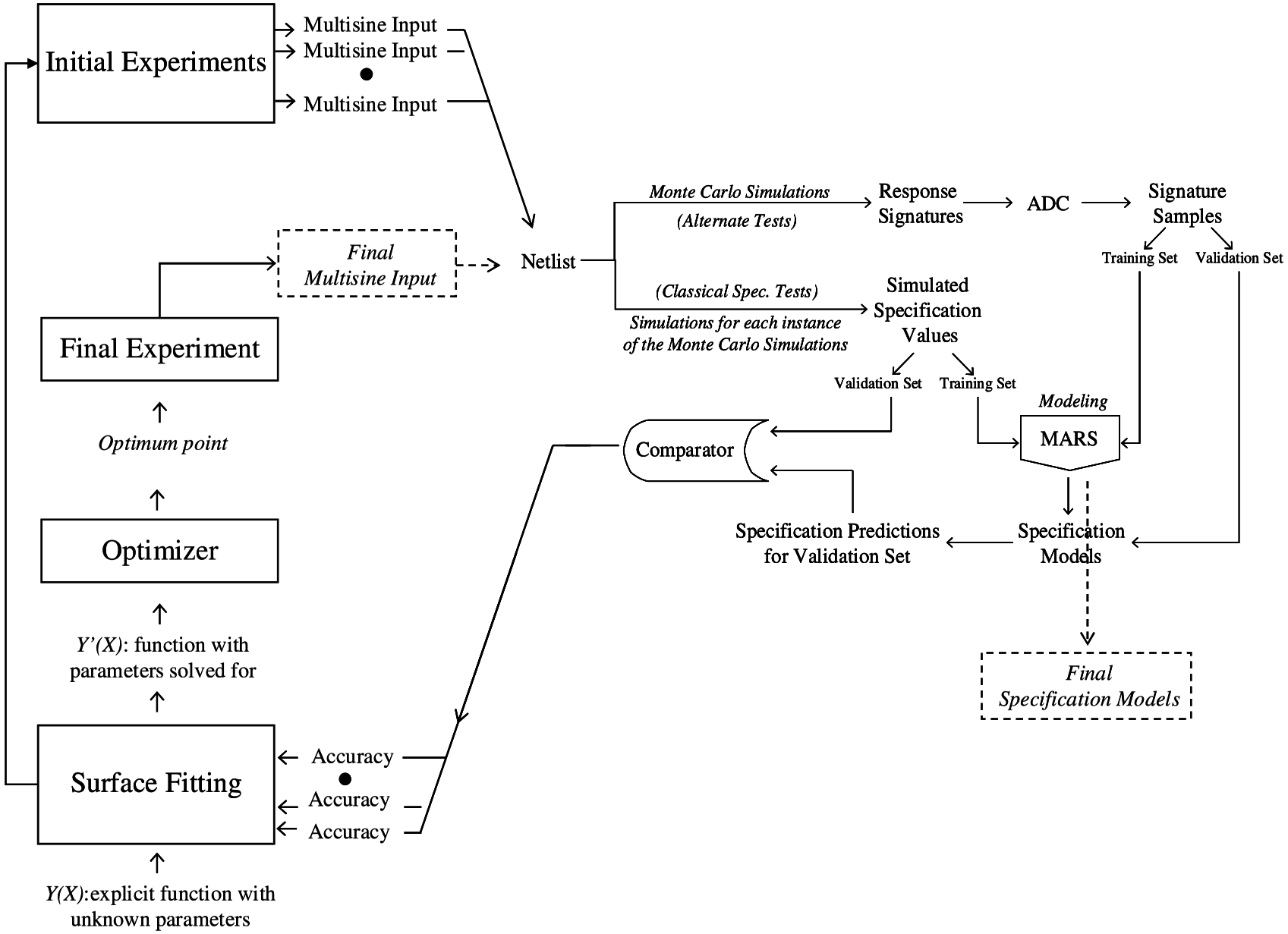


Figure 13: Implementation of an analytical optimization solution.

Another solution to optimization of explicit relations is formulated by the general class of mathematical programming. Mathematical programming methods systematically seek optimal solutions by means of computational procedures. Two popular methods applicable to linear functions are linear programming and goal programming. In linear programming, both the objective function, explicit relation, and the constraints are formalized as linear combinations of independent variables. Goal programming extends this approach by providing solutions in the presence of multiple objective functions. These methodologies can be extended to the nonlinear domain by solving the linearized forms of the original problem. Such methods include quadratic programming, separable programming, approximation programming, and cutting plane procedures. The methods of feasible directions address the same problem without the need for linearization, but require a feasible starting point. The most effective form of nonlinear programming is the class of methodologies known as penalty functions, in which constraints are integrated to the object functions in the form of penalty functions. Other nonlinear programming techniques include geometric programming, dynamic programming, and nonlinear goal programming.

Both the analytical optimization and linear/nonlinear programming methodologies handle the optimization problem separately from the experimentation. However, if the problem itself is not already formulated, it will be the experimentation procedure that defines the objective and the constraint functions necessary for the optimization. In this sense, the problem may be considered as a black box where the relation between the response and the independent variables is not known in advance. The procedure starts from some initial values for the independent variables and performs a set of experiments to produce the corresponding response variables. The optimization phase uses these experimental responses to provide the values of independent variables to be used in the next set of experiments. Following this "optimization through experimentation" cycle, the independent variables converge to an

optimal set. Several methodologies have been presented in the literature as a solution to this formulation. One such solution is the class of response surface methods. Response surface methods model the unknown relationship between the response and the independent variables through designed experiments.

A simple but effective implementation of response surface methodology is an iterative modeling procedure. The procedure starts with performing a number of experimental trials, and then fits linear models for the response variables. The gradient of the resulting model is solved for an optimum point, and the next set of experiments is designed around this point. The search continues until the optimum point along the last gradient is close enough to be the global optimum. One can apply sophisticated search techniques to solve for the optimum point along the gradient. Such well known techniques are the sequential simplex method, the convex method, and the pattern search. Mathematical programming can also be integrated into response surface methods. The procedure starts with a large number of experiments over an initial region and maps the observations into quadratic models. Then, these models can be used as the objective functions for the nonlinear programming procedures together with the constraints on the origins of the initial region. The resulting optimal solution is used as the basis for the new set of experiments.

Figure 14 shows the integration of response surface models into the specification-based alternate test methodology. First, the controlled experimentation module generates a number of multisine inputs. Second, each input goes through the simulator-based process, generating a prediction accuracy. Then, all the accuracy values are used to produce models for each one of the specifications, and the optimization procedure solves for the optimum value dictated by these models. Finally, the resulting optimum is fed back into the controlled experimentation module. The optimization through experimentation cycle runs until the prediction accuracy satisfies the test coverage requirements.

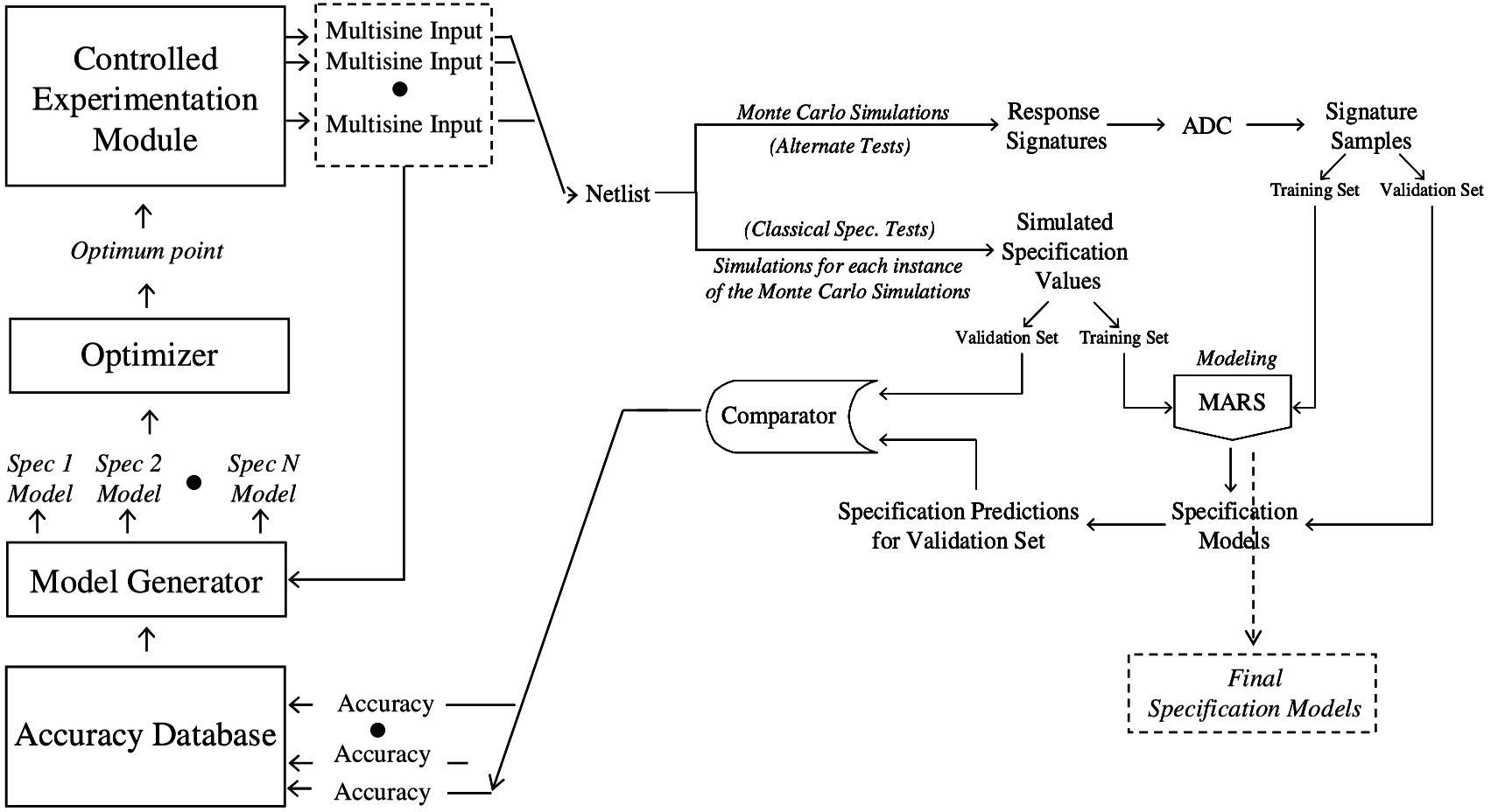


Figure 14: Response surface model solution to proposed methodology.

The algorithm for the selection of sinusoidal parameters can be summarized in four steps as shown in Figure 15: Every iteration starts with the selection of an initial vector of n superimposed test frequencies. The next step computes the gradient of the cost function with regard to the amplitudes and frequencies of the sinusoidal waveforms. Then, using the gradient vector, the controlled experimentation module picks the next choice of test frequencies. The module may also merge or split the choice of test frequencies. If two individual tones move close enough after a gradient operation, they are merged. If a tone ends up on a peak of the gradient space, it is split and both directions are investigated. Figure 16 depicts an example of five iterations showing split and merge operations. The algorithm stops if no further improvement in cost is possible; otherwise execution proceeds from the first step with the selection of a new vector of test frequencies. The experiment controller. which assigns new *(w,a)* pairs and splits or merges frequencies, acts in favor of exploration if recently no major improvement is recorded in cost (c). It acts in favor of exploitation if recently gradient vector (v) yields major improvement in cost.

3.1.4 Implementation Example 1: LNA

To demonstrate the algorithm, the procedure is applied to a 900MHz LNA, as described in [49]. Figure 17 shows the schematic of the LNA with 8 resistors, 5 capacitors, and 2 transistors. The saturation current and the forward gain of the transistors, together with resistor and capacitor values sum up to 17 process variables. Each process variable is assumed to have a normal distribution with

$$3\sigma = \frac{nom}{10}, \tag{3}$$

where nom is the nominal value for the variable, and σ is the standard deviation. The validation set is a 500-sample random set with a jointly normal distribution. The training set is composed of two parts, the first one is a 500-sample jointly normal

0. *Initialize variables*
1. *compute cost at the new point*
2. *if cost > minimum desirable or iteration > maximum number of iterations go to 14*
3. *if |cost – previous cost| < minimum cost change go to 12*
4. *iteration <- iteration + 1*
5. *compute the gradient function v of frequencies w_i and amplitudes a_i wrt cost function c*
6. *compute the new frequency and amplitude values using the gradient*
   move in the steepest descent direction of cost
7. *for any one of the frequency elements in v that is close to zero do*
   i. *compute second order partial derivative wrt the element*
   ii. *if second order partial derivative greater than 0*
   local minimum => fix the element value
   iii. *else local maximum*
   a. *if number of frequencies < maximum number of frequencies*
   split the frequency into two frequencies, regulate amplitudes
   b. *else avoid local maximum by stepping to an arbitrary direction*
8. *for every w_i, w_j pair*
   if the distance between w_i and w_j < minimum frequency distance
   merge w_i and w_j ; merge their amplitudes
9. *save frequency, amplitude, gradients, and cost data with the index value of iteration*
10. *update new frequency and amplitude values based on changes in 6, 7, and 8*
11. *go to 1*
12. *propose a new set of frequency and amplitude values*
13. *go to 1*
14. *select the best cost data out of all iterations*
15. *output the frequency and amplitude values corresponding to the best cost data*

$$\begin{bmatrix} w \\ a \end{bmatrix} = \begin{bmatrix} \omega_1 & \omega_2 & \dots & \omega_n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{bmatrix} \qquad v = \begin{bmatrix} \frac{\partial c}{\partial \omega_1} & \frac{\partial c}{\partial \omega_2} & \dots & \frac{\partial c}{\partial \omega_n} \\ \frac{\partial c}{\partial \alpha_1} & \frac{\partial c}{\partial \alpha_2} & \dots & \frac{\partial c}{\partial \alpha_n} \end{bmatrix}$$

Figure 15: Tone selection algorithm.

distribution; whereas the second part is another 500-sample random set with process variables linearly distributed over the $\pm 10\%$ range around the nominal values as shown in Figure 18. This combination of normal and linear sets provides a large coverage of possible faults, yet preserves the nature of a realistic distribution. The validation and training set instances are generated by the MC method.

The specifications of interest for the LNA are chosen such that each one emphasizes a different aspect of transient testing. The sample specifications are 1dBComp, IIP3, and NF at the nominal operating frequency. The 1dBComp is a good figure of performance for single tone inputs, whereas IIP3 is typically measured by two-tone

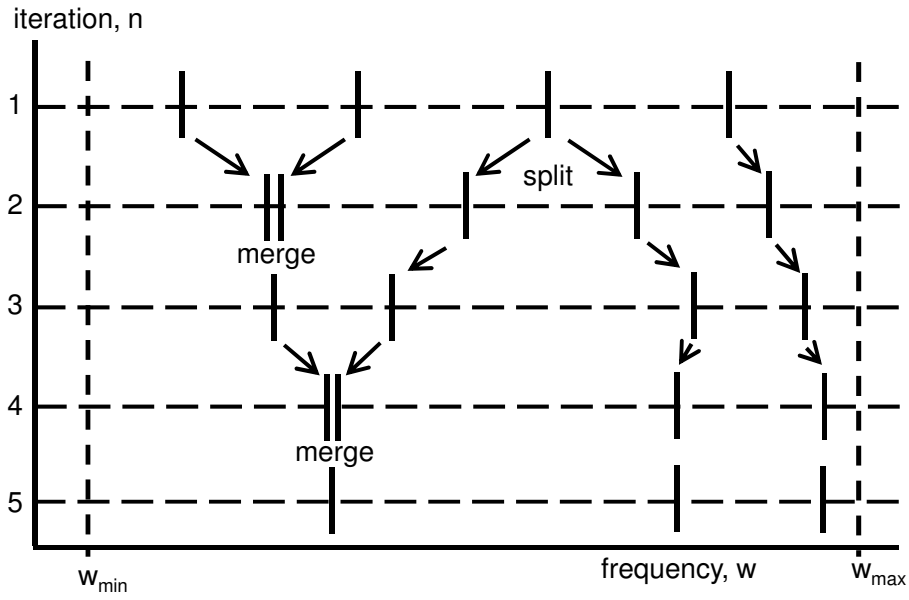


Figure 16: Split and merge operations.

inputs. Noise figure is highly frequency dependent. Table 3.1.4 summarizes the nominal, minimum and maximum values for the distribution of these specifications and the corresponding simulation methods.

Table 1: Specification distribution and analyses types for the LNA example.

Specification	Analyses	Nominal Value	Minimum	Maximum
1dBComp	pss	-10.6	-13.0	-8.94
IIP3	pss	1.82	0.0234	3.96
NF@900MHz	pnoise	4.100	4.082	4.122

The optimizer can make use of any prediction accuracy metric, that is well-behaved mathematically. The computation of two of these metrics, the maximum percentage error *(mpe)*, and the root mean square percentage error *(rmspe)* are given in equations 4 and 5, where "actual" is the actual value of the specification, "predicted" is the value predicted by MARS, and "rms" is the root-mean-square function. The maximum and root-mean-square functions are computed over the validation set.

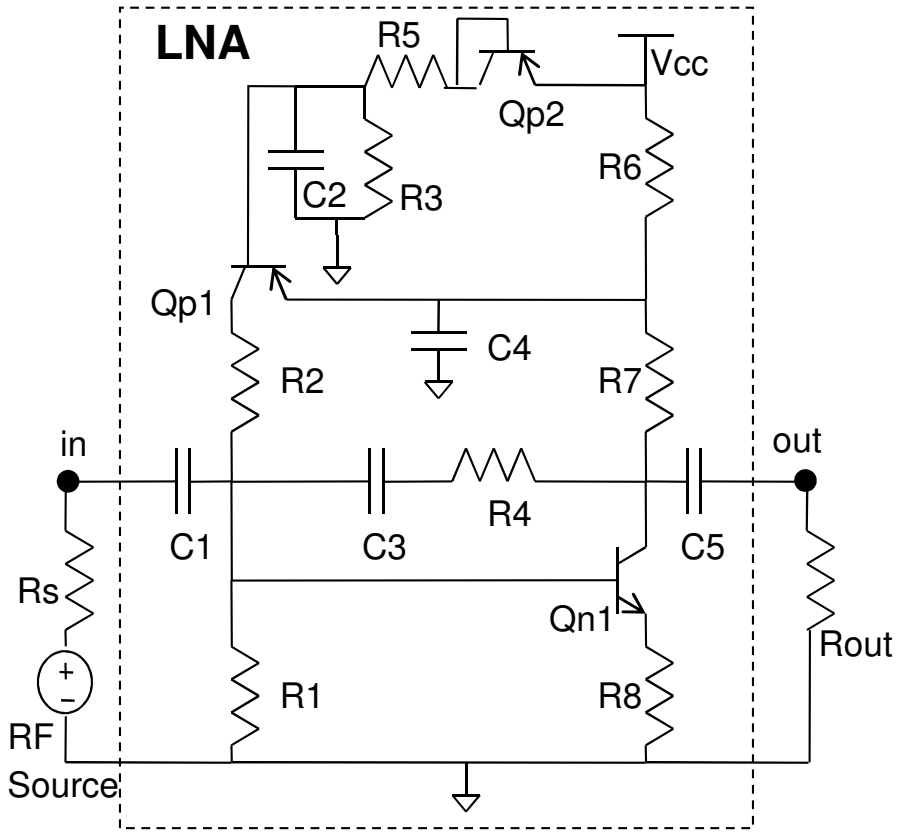


Figure 17: The schematic of the 900MHz LNA.

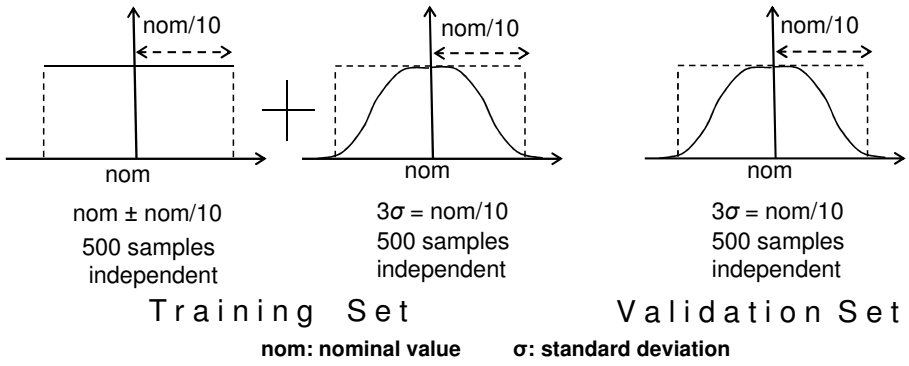


Figure 18: Process parameter distributions for training and validation sets.

$$mpe = maximum\left(\frac{|actual - predicted|}{actual}\right) \tag{4}$$

$$rmspe = rms\left(\frac{|actual - predicted|}{actual}\right) \tag{5}$$

Mpe is a critical constraint in evaluating the success of predictions. However, due to the nature of the regression algorithm, maximum values of error localize either around the extremums or right around the mean values. Any of these two regions rarely lead to pass/fail misclassifications provided that the prediction error is within some reasonable limit. Experiments show that an error measure collecting the behavior of the overall specification domain, such as *rmspe*, better fits as an iterative optimization metric. The main instances of interest are the ones that have specifications close to the pass-fail thresholds. Neither *mpe* nor *rmpse* specifically gives an insight on the prediction accuracy around these regions. In order to handle this issue, we introduced quantitative metrics of misclassification into the algorithm. Two metrics quantify the instances that are false positives, instances that should fail but classified as passing by prediction, and the ones that are false negatives, instances that shoud pass but classified as failing by prediction. In this experiment, the pass/fail tolerances for each specification are arranged such that 5% of the instances in the validation set fail, for lack of a better pass/fail criteria. NF error tolerance is the smallest hence the most critical one, whereas IIP3 error tolerance is distributed over a wide range, which means that the *mpe* metric has a rather weak correlation with classification metrics.

Some selected solutions for the MUSTI algorithm are presented in Table 3.1.4. All parameters for these solutions are the same except for the number of tones and the frequency values of the tones. The first row describes the classical test for IIP3; one of the sinusoids is the nominal operating frequency of 900MHz, and the other is superposed at 920MHz. Although the maximum error in IIP3 prediction is close to 10%, none of the instances are misclassified in terms of IIP3. Overall, there is only 1 misclassified instance, which corresponds to 0.2%. The solution in the second row replaces the two-tone input with a single tone at the nominal operating frequency. The corresponding prediction accuracy and error distributions are shown in Figure 19.

Maximum percentage error metrics are better in this case, pulling IIP3 error below 1.5%. The percentage of misclassified instances increases only by 0.4%. At an even lower frequency of 868MHz, the misclassification is as low as that of a two-tone sinusoidal input, and the maximum percentage errors get smaller. This is a remarkable point in the sense that our algorithm generates the same fault coverage as a two-tone signal using only a single tone, the frequency of which is lower than the nominal operating frequency, and also provides lower maximum error values. The fourth solution performs almost as good as the second one at a frequency value less than half of the nominal. This is also significant in the sense that, our algorithm cuts the maximum frequency requirement of the ATE by more than 50%, if a 0.6% misclassification is acceptable. These results show that there is a trade off between the specification coverage and minimum possible frequency value. The fifth solution shows that at a frequency two orders smaller than the nominal operation, the correct classification percentage for 1dBComp and IIP3 are still high, whereas a misclassification of 2.0% is seen for NF.

Table 2: Selected solutions for the LNA with MUSTI algorithm.

Tones	Freq.	Max % Error			# of False Neg.			# of False Pos.			Misclass.
#	(MHz)	1dB	IIP3	NF	1dB	IIP3	NF	1dB	IIP3	NF	# (%)
2	900&920	1.44	9.75	0.04	0	0	0	0	0	1	1 (0.2%)
1	900	0.84	1.40	0.08	0	0	2	1	0	1	3 (0.6%)
1	868	0.56	1.41	0.05	0	0	0	1	0	0	1 (0.2%)
1	411	0.42	1.69	0.06	0	0	1	1	0	1	3 (0.6%)
1	7	9.57	5.38	0.17	1	0	2	0	0	3	6 (2.0%)

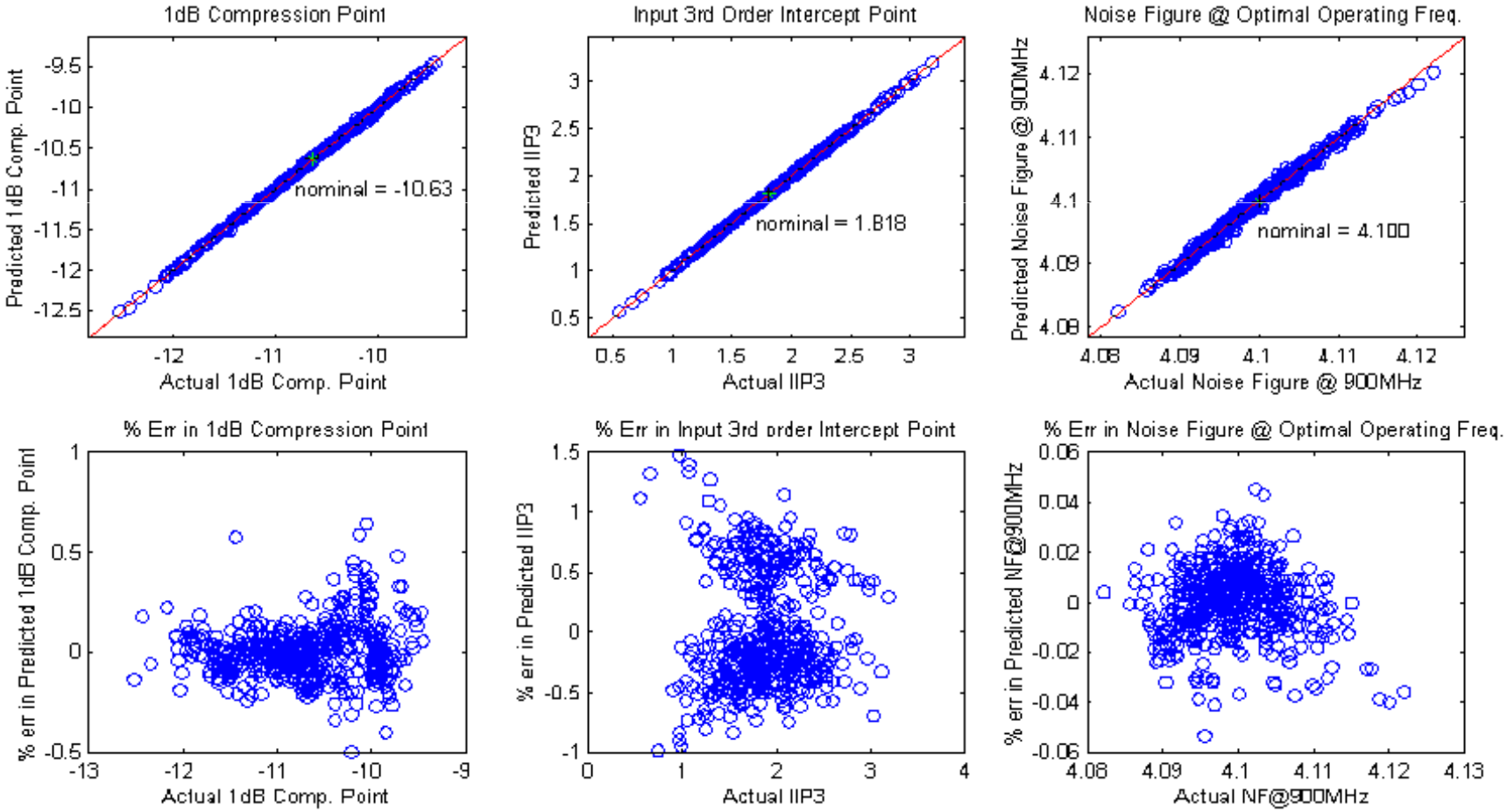


Figure 19: Actual vs predicted specifications of the LNA using MUSTI algorithm with single-tone stimulus at 868MHz.

Figure 20 shows the effect of number of ADC bits on the prediction accuracy. All the parameters except ADC bits are kept constant for the 868MHz sinusoidal input. The graph shows that the performance decrease significantly with decreasing number of bits when the number of bits is smaller than 7. When it is above 12, it effects the prediction accuracy in a minor way. The prediction accuracy falls below 2% for bit numbers larger than 6, and it is smaller than 1% for 11 bits and above.

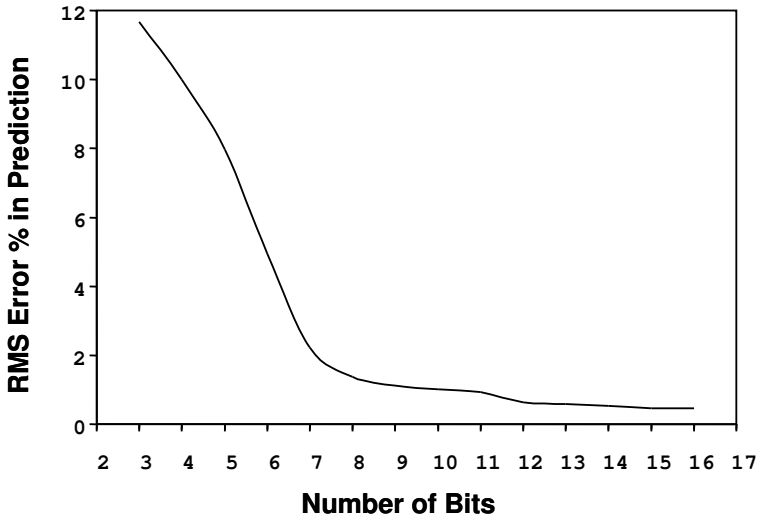


Figure 20: ADC bits versus prediction accuracy for the 868MHz single-tone input.

Figure 21 shows the effect of number of samples on prediction accuracy. The same model generation routine is applied for different number of samples of the 868MHz single-tone response. The graph shows that it is possible to reduce the number of samples down to 10 with little sacrifice in prediction accuracy.

3.1.5 Implementation Example 2: Mixer

The previous example showed that MUSTI is a suitable and efficient way to produce specification based alternate tests that run at speeds much lower than the nominal operating frequency. On the other hand, even for single-tone tests, the maximum error is a low value on the order of 1.5% to 2%, which makes it very difficult to

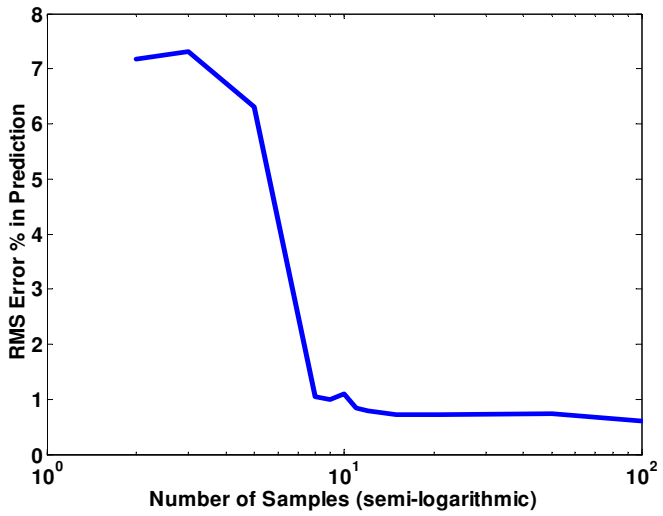


Figure 21: Number of samples versus prediction accuracy for the 868MHz single-tone input.

find multitone tests that can compete with single-tone tests. The iterations of the algorithm that start with multitones either converge to sub-optimal solutions or all the frequency values are simply merged into a single-tone.

In order to demonstrate the multitone solutions of MUSTI algorithm, a second example is studied in this section. Figure 22 shows the schematic of a 920MHz down-converter mixer driven by a 1GHz local oscillator (LO). The saturation current and the forward gain of the transistors, together with the sheet resistance sum up to five process variables. The sample specifications of interest for the mixer are 1dBComp, IIP3, conversion gain at 920 MHz (CG@920MHz) and power supply rejection ratio at 920 MHz (PSRR@920MHz). Table 3.1.5 summarizes the nominal, minimum, and maximum values for the distributions of these specifications and corresponding simulation methods.

Table 3.1.5 shows multitone solutions for the mixer. The single-tone solution starts with three tones covering a spectrum around the nominal operating frequency of the mixer. The two-tone solution also starts with three tones, but they cover a wider

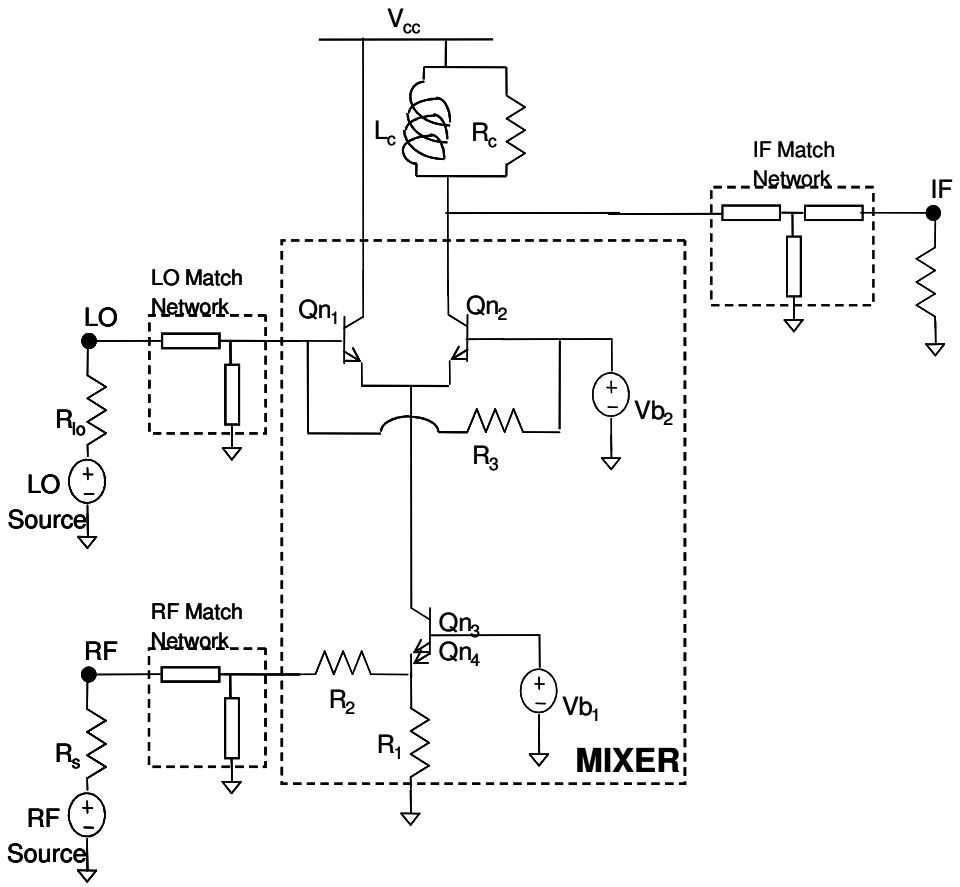


Figure 22: RF downconversion mixer.

spectrum. The three-tone and four-tone solutions are optimizations of four-tone and six-tone initial points. Figure 23 shows the prediction accuracy and error distributions for the four-tone solution.

Table 3: Nominal, minimum and maximum values for the distribution of mixer specifications, and corresponding types of analysis.

Specification	Analyses	Nominal Value	Minimum	Maximum
1dBComp	pss	-4.87	-5.13	-4.50
IIP3	pss	8.15	7.42	9.07
CG@920MHz	ptf	-5.44	-5.64	-5.21
PSRR@920MHz	ptf	-73.1	-74.2	-71.7

Table 4: Selected solutions for the mixer with MUSTI algorithm.

# of Tones	Starts with	Freq. values (MHz)	Max % Error			
			IIP3	1dB	CG	PSRR
1	3 tones	660	1.0	1.5	0.06	0.08
2	3 tones	670,440	1.1	1.2	0.01	0.1
3	4 tones	720,390,160	1.1	0.8	0.1	0.07
4	6 tones	860,540,370,210	0.9	0.6	0.04	0.08

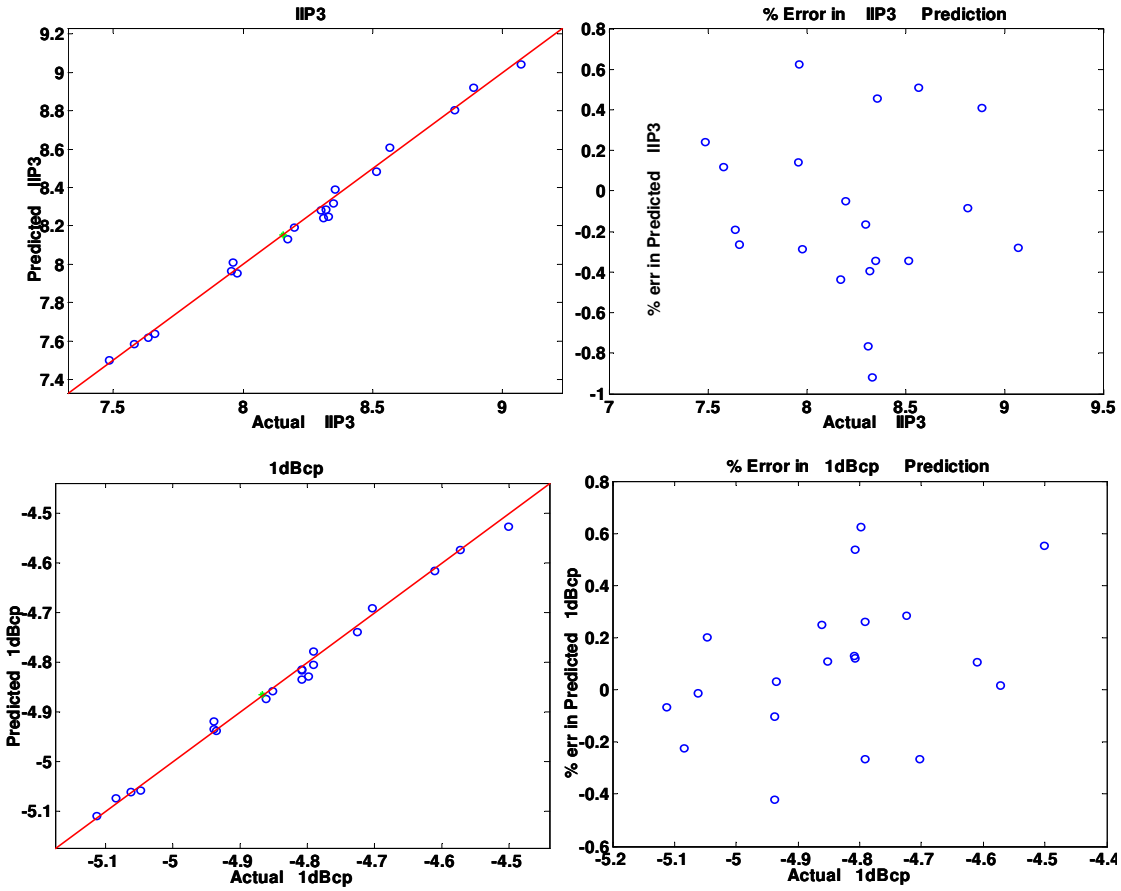


Figure 23: Four-tone test results for RF downconversion mixer: IIP3 & 1dBComp.

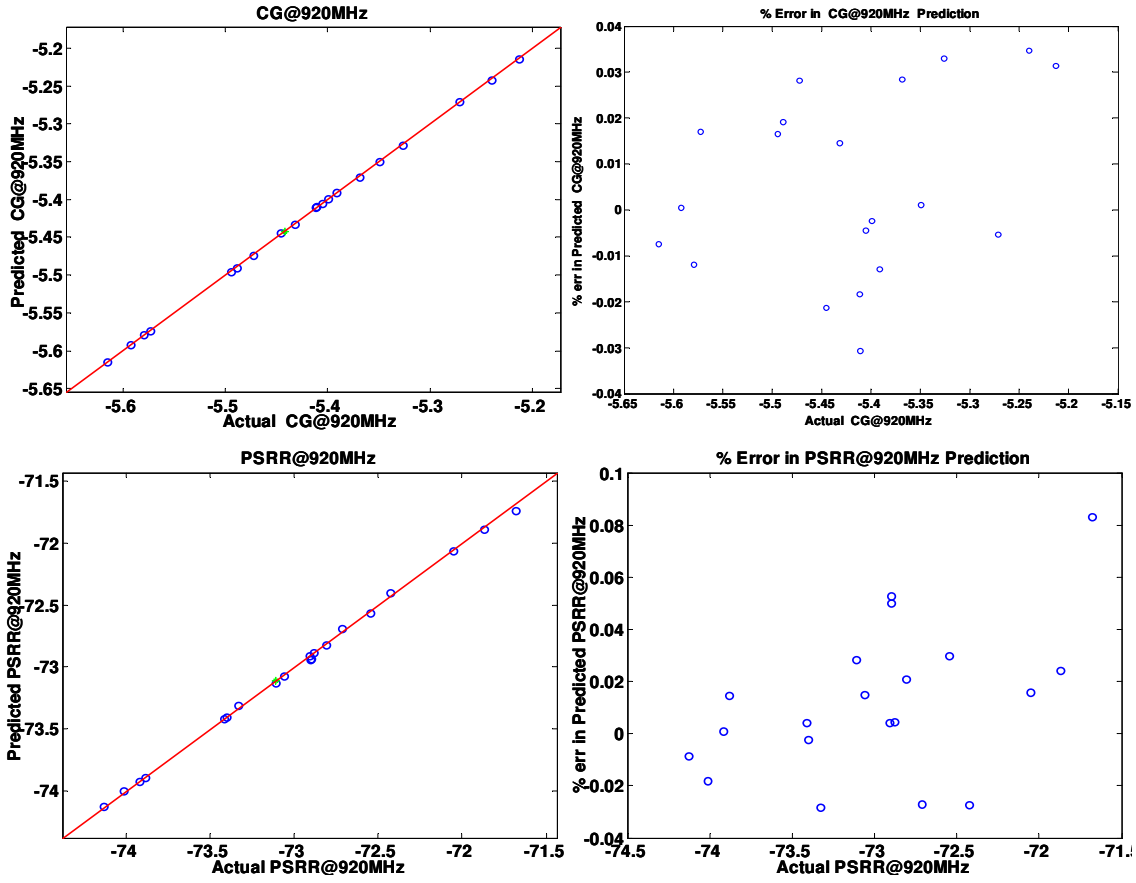


Figure 23: Four-tone test results for RF downconversion mixer (cont'd): CG & PSSR.

3.2 Test Stimulus Generation with Current Signatures

In this section, a technique for low-cost testing of RF components integrating current signatures and alternate test methodology is presented. The technique is suitable for non-invasive BIT as well as low-cost ATE applications. Main features of the technique are (i) minimum loading on signal path by sampling supply current, (ii) flexible test stimulus generation based on system constraints, (iii) test time reduction by using a single test stimulus and data acquisition, and (iv) accurate prediction of all specification values from the single excitation. Two experiments using the proposed implementation demonstrate the accuracy and efficiency of the technique on both single-balanced and double-balanced mixers built with two different technologies.

3.2.1 Motivation and Background

To reduce tester costs and limit the test bandwidth requirements, there is a trend to move tester functionality to the close proximity of the DUT in the form of BIT and BOT. BOT migrates high-speed functions of the ATE to the tester loadboard, while BIT implements these functions within the same chip or package. In this section, we propose a technique for low-cost testing of RF components with supply current signatures. The technique is suitable for BIT applications with on-chip/package ADC and DSP. When additional resources are not available on-chip/package, the technique can be implemented as a BOT scheme for all-digital ATE with ADC and stimulus generator on the loadboard. The automated stimulus generation and indirect specification measurement is handled by the alternate test methodology.

This work addresses the following challenges in BIT and BOT for RF components: (i) BIT/BOT area overhead and increased test time resulting from dedicated measurement features, (ii) parasitics and power loss resulting from BIT components on the signal path, (iii) constraint-driven stimulus generation for BIT, and (iv) auto-calibration

of process variations on BIT circuitry. All high-speed features necessary to test different SUTs cannot be practically implemented on the loadboard or on-chip. The practical limitation does not come from circuit design but rather from the economics of area overhead in BIT and loadboard complexity in BOT. The proposed direct measurement techniques using built-in/off features [57–59] are dedicated to a single type of specification test and will require a significant amount of area overhead -up to 15%- when individual solutions are integrated. Recently, some solutions are presented where closely related specifications are directly measured by a single BIT resource: [60] combines gain and 1dBComp, whereas [61] also computes IIP3 by mathematically deriving it from the other two. Both solutions target component level specifications but they are limited to gain and compression related specifications, whereas isolation, NF and port mismatch needs to be handled by extra built-in circuitry. A system level BIT approach is demonstrated in [62] which makes use of power detectors and frequency response characterization circuitry in loopback mode. Although it provides 4.8% area overhead for a large digital enhanced cordless telecommunications (DECT) transceiver -still excluding noise related specification measurements such as signal-to-noise ratio SNR), spurious-free dynamic range (SFDR) or bit error rate (BER)-, the overhead becomes 7.6% for a smaller Bluetooth transceiver; hence the extra area will be prohibitive for BIT of stand alone components such as mixers.

The ideal solution is a general one which will make use of a limited number of BOT/BIT resources to determine all specifications of interest and can scale with the complexity of the DUT from simple components to complex systems. Alternate test methodology is a suitable candidate, which replaces a number of ordinary tests with a single test performed by less number of resources. Advantages of alternate tests have been demonstrated in specification-based testing and fault diagnosis of analog circuits [42,63]. In addition to reducing area overhead, this approach significantly cuts down the testing time. A detailed comparison of dedicated BIT/BOT techniques and

generalized alternate test solutions is presented in [39]. Alternate tests are designed such that the response signature of the DUT is highly correlated to the SUTs in the presence of process variations. Other than simple pass/fail decisions, the test results can also be in the form of quantitative specification value predictions so as to provide a measure for the quality of each prediction. Built-in applications and integration feasibility of the methodology is featured in [39, 45, 46, 64–66].

The parasitics associated with test electronics constitute one of the major problems for BIT of RF components, which are carefully designed for matched loads and cannot tolerate additional power loss associated with test multiplexing and coupling circuitry. These negative effects have to be compensated at the design phase of the DUT, or otherwise there is a severe impact on the device performance. Monitoring supply current proposes a non-invasive alternative to such problems. Since the observation node is not on the critical signal path, supply current can be copied and monitored without significant degradation of circuit performance.

Another challenge for BIT/BOT is stimulus generation, the on-chip/board features may be limited in frequency and power. One solution is making use of an external low-frequency test input and upconverting it to the RF range using built-in/off electronics; similarly a downconverter can be used for sampling at lower signal rates [43]. Alternatively, one can use lower frequency signals to stimulate the DUT and sample the response to predict specifications at the higher operating frequency, as was shown in the previous section. In both cases, the accuracy of the specification prediction heavily depends on stimulus power. In some cases, the correlation between the specifications and response signatures are strong only when the stimulus has high power levels, which are not practical for BIT applications. This is especially true when some of the specifications under test involve nonlinear characteristics apparent only through power compression, such as 1dBComp or IIP3. In this work, the first

experiment shows that the specification prediction accuracy is not significantly degraded by decreasing stimulus power, when the response signature is generated from the supply current as opposed to output voltage.

Another major challenge is the problem of testing the BIT circuitry itself. Since the fault mechanisms, process variations and environment variables such as temperature will affect both the DUT and the BIT circuitry, the BIT circuits need to be tested beforehand or they should be able to adapt and compensate for the expected variations. In [66], the authors suggest that a number of BIT detectors scattered around the system will form a redundant network that will raise a flag when the BIT responses are inconsistent. This way, the problems with BIST circuitry can be detected but cannot be compensated. Typical solutions are running a calibration scheme before the actual test [60], or making BIT immune to variations in a certain range with additional circuitry [67]. These solutions present trade-offs between accuracy, test time, and area overhead. Another solution makes use of much simpler BIT circuitry, which mimics common structures already present in the DUT [46]. This way, the effects of process variations are highly correlated in the DUT and the BIT circuitry, hence machine learning techniques that are inherent in alternate test can auto-compensate the deviations in BIT performance. In this work, we follow the idea in [46] and study such effects with the second implementation example.

3.2.2 Current Signatures in Testing

Supply current-based testing has extensive applications in digital systems. *Iddq* testing provides a powerful and cheap tool for wafer-level test of digital systems built of MOSFETs, and its extensions have the potential to eliminate burn-in test [68]. Although submicron devices proposed a challenge with increasing leakage currents, variations of *Iddq*, such as *delta Iddq*, enhance the resolution [69], partially eliminating the problem. A generalized theory of current signature analysis and its extension

to transient supply current monitoring is described in [70]. Finally, [71] presents a built-in architecture for the on-line monitoring of quiescent and transient supply currents, while [72] demonstrates a similar built-in implementation for analog structures. An RF example is demonstrated in [73].

Although the use of current signatures proposes advantages, it also suffers from disadvantages specific to RF testing. For example, it may not be suitable for all kinds of RF components, especially for ones that rely heavily on bias currents. In such cases, the supply current may be dominated by the bias component and the variation resulting from parametric variations may be below the signature resolution. A typical example is LNAs made of bipolar devices. Experiments on the 900MHz LNA shown in Figure 24 focus on the supply current variation on a Monte Carlo set of 400 devices. The sample set was subject to 30% variation in transistor parameters and resistance values. Although SUTs (gain, IIP3, NF, 1dBComp, reverse isolation RIso, input-output matching Sin and Sout) show significant variations (3.6dB, 8.2dB, 1.7dB, 9.8dB, 1.4dB, 1.1dB, and 1.3dB respectively), the maximum variation in supply current is only 2.1% with a standard deviation of 0.62%. A similar experiment is performed with a 2.4GHz CMOS LNA [47] designed in 0.25μ technology as shown in Figure 25. For significant distributions in specifications, the supply current shows only 2.6% maximum variation. These experiments show that test schemes based on supply current measurements may not be suitable for LNAs with significant bias currents.

Power amplifiers (PAs) suffer from the similar problem of high bias currents. One recent alternate test paper [73] applies a ramp signal as the bias current, and uses time-domain supply current samples as the DUT signature. This way the strong correlation to the bias current is taken advantage of, however this approach is limited to devices with external bias control and difficult to implement as a BIT solution.

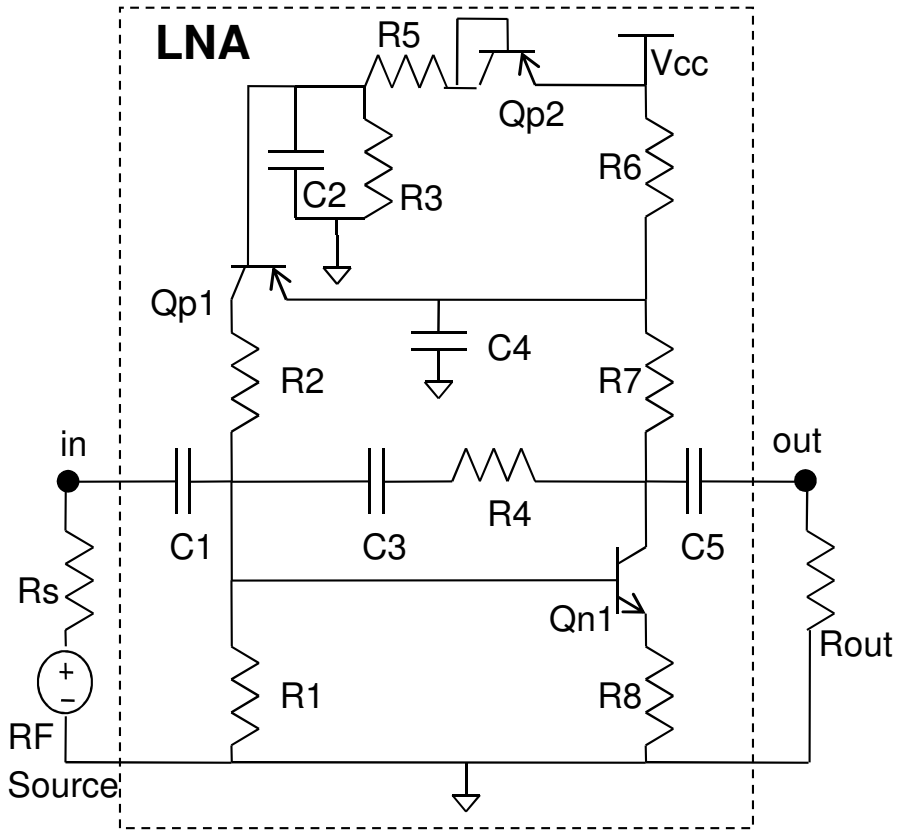


Figure 24: The schematic of the 900MHz bipolar LNA.

Improvement can be achieved if currents of multiple stages of an amplifier can be monitored separately, this is partly possible for the bipolar LNA example, whereas the CMOS LNA has only one power path. When different branches of supply currents can be isolated and individually sampled, current-based test can provide better observability by revealing faults/variations in earlier stages, which may otherwise be masked by following stages of the circuit. While voltage-based sampling will create loading problems and degrade the signal from one stage to another, current-based solutions will only require additional area.

Another solution to high bias current is observing the difference between the supply current of the DUT and the ideal bias current of the design. Although similar *delta Iddq* approaches have been taken in the digital domain, this is considered as an off-chip solution. A current subtractor can be implemented on-chip as shown in

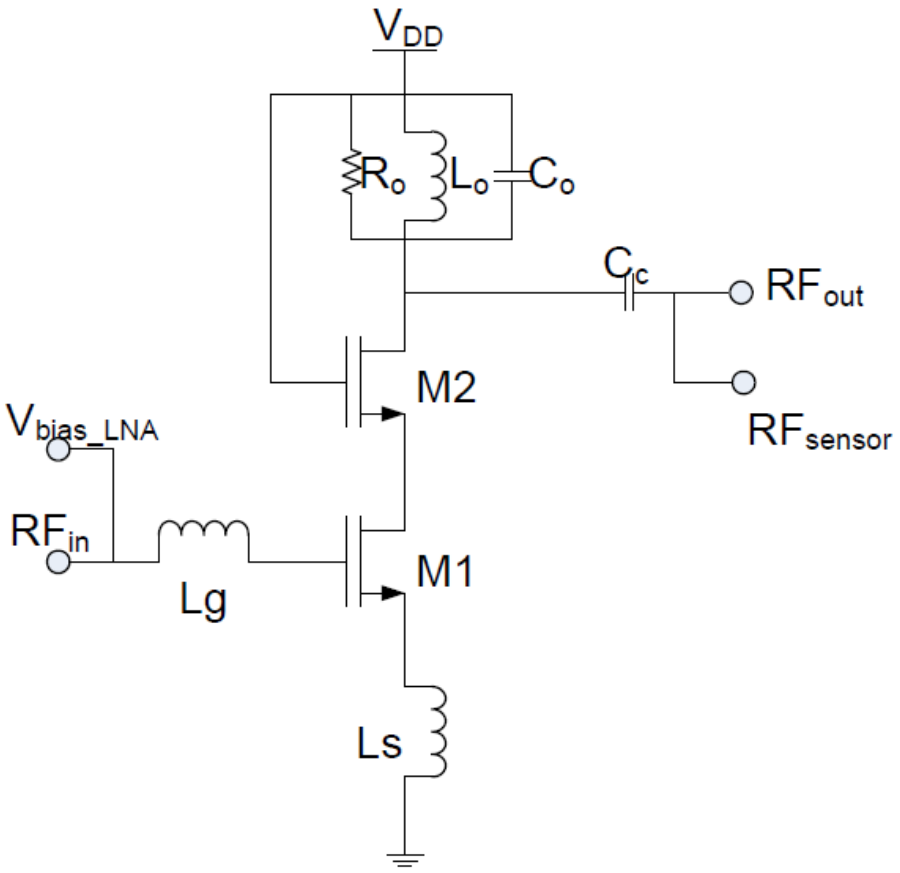


Figure 25: 2.4GHz low-noise amplifier.

Figure 26: a current mirror copies the supply current, *Idut*, to the branch on the right, the difference between the DUT current and the nominal bias current, *Inom*, provided by transistor *M3* is fed into the I-to-V converter as the DUT signature. Since this configuration requires a complete copy of the bias current, it doubles the power consumption. Although transistor ratios can be arranged so that a fraction of the DUT current is copied, this will also scale down the final signature.

Additional copies of DUT current are an important problem for PAs. Such BIT schemes are not practical since it will double power consumption, threaten the reliability of the package and cause temperature drifts. Instead of mirroring the current, one can amplify the voltage across a sense resistor on the power path. This scheme

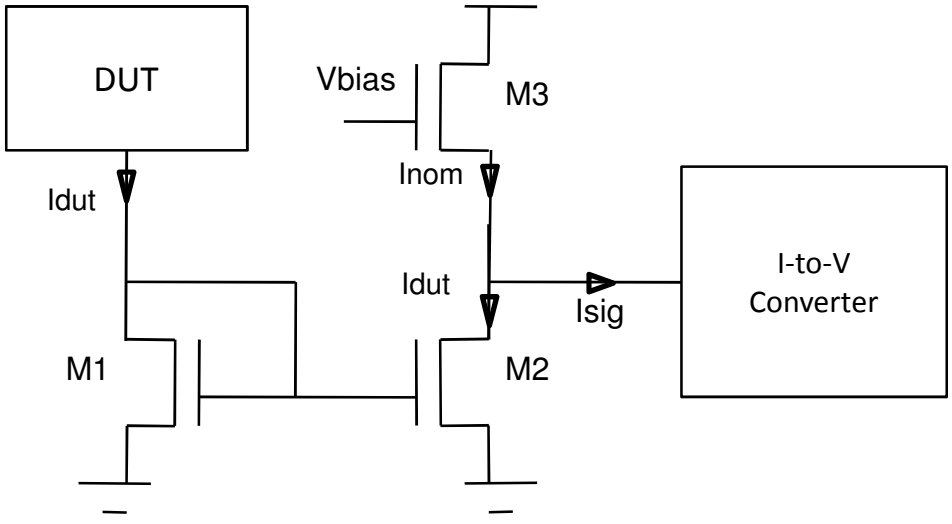


Figure 26: Current mirror configuration for *delta-Iddq*.

has been applied to LNAs in [74, 75]. Although the proposed current sensor has very low area overhead (0.1%), it requires far more active and passive elements than the LNA itself. It is not clear how this complexity affects the reliability of the sensor under process variations. In this work, a similar sense resistor is used, but instead the current sensor is kept as simple as possible so that variations in the DUT dictate similar variations in the test circuitry. Machine learning process in alternate test tracks these changes and automatically compensates for variations in the BIT circuitry.

3.2.3 Methodology

The DUT is excited by an alternate test stimulus, with the general test generation scheme shown in Figure 27. First, a Monte Carlo (MC) set of DUT instances created with given variations in process parameters. Each instance is simulated for target specifications. Then, in each iteration of the optimization loop, candidate stimuli are applied to the MC set. After the circuits are stabilized, the supply currents are monitored, filtered and sampled into digital signatures. They are fed into the specification mapping functions producing the predictions for each SUT. They are

compared with the actual specification values and the stimuli are tweaked depending on the accuracy of the predictions.

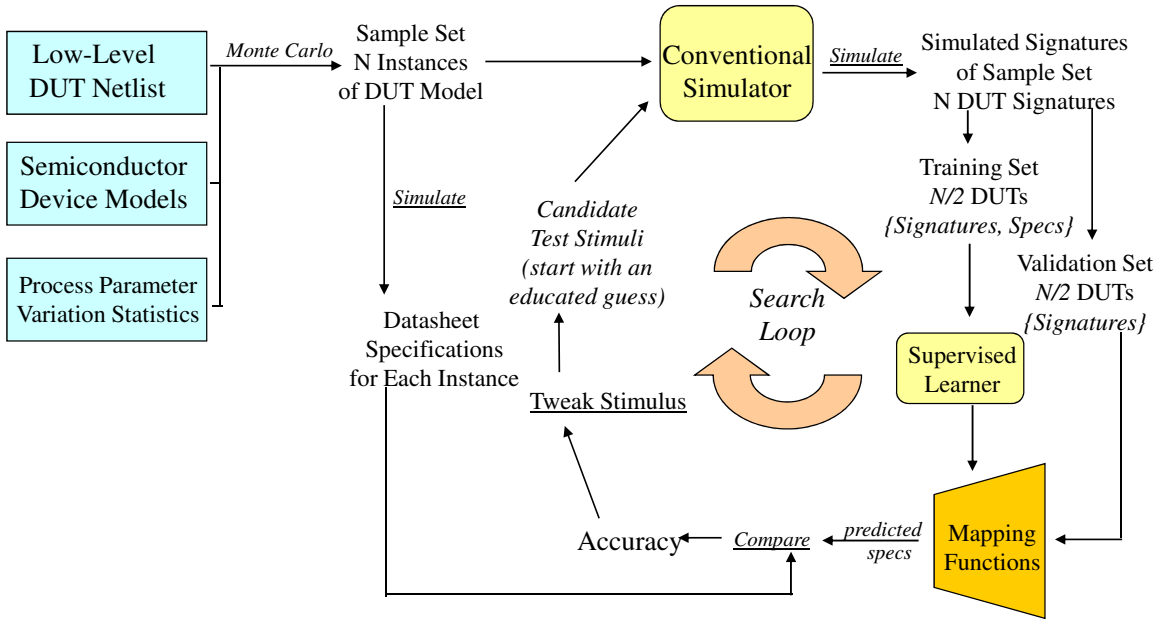


Figure 27: Test generation scheme for alternate test.

In the first example, we will search for a single-tone stimulus, the simplest configuration for an on chip/package stimulus generator. The goal of the test generation is to find the frequency and amplitude of the sinusoid that best exploits the correlation between the current signature and the SUT. Since the search space is relatively small, we used a greedy algorithm as described in the previous section with digitized transient samples as the corresponding signatures.

In the second example, we used the genetic algorithm (GA) based optimization scheme that will be studied in detail in the next section. The number of possible tones at the input is limited to two, since more tones will be harder to generate on chip/package. As the number of tones is limited, the GA algorithm can use transient simulations and netlist-level models without a large computational penalty. The amplitudes of the dominant FFT components are the corresponding signatures in this case.

3.2.4 Implementation Examples

The proposed methodology is demonstrated with two mixer examples. The first example is a proof of concept, in which it is assumed that the supply current is properly sampled by a current sensor. The feasibility of the proposed methodology is checked by imposing two constraints on the alternate test generator: (i) single-tone test stimulus, and (ii) small stimulus amplitude. Also the current-based results are compared with voltage sampling. In the second example, the current sensor is implemented and its effects on the original mixer specifications are quantified together with its ability to automatically calibrate for possible process variations.

Figure 28 shows the schematic of the first example, 920MHz downconverter mixer with a 1GHz LO. This is a single-balanced commercial mixer built with bipolar transistors, the design is available from the Cadence Spectre toolkit. The saturation currents and forward gains of two different kinds of transistors, together with the sheet resistance are considered as process variations. 150 MC instances are generated for the training set, and 50 independent instances are used to validate the mapping. Although the output of a downconversion mixer does not exhibit significant frequency components above the downconversion frequency, 80MHz in this case, its supply current reflects the effects of LO, around 1GHz, and has much more complex frequency characteristics. In this form, the current waveform presents a rich signature and produces accurate mappings. However, in order to be able to compare its performance with that of voltage based mapping, two waveforms have to be sampled at the same rate. This process requires an analog lowpass filter before the sampling step. For the sake of simulation time, the filter is created in digital domain using Matlab. The analog filtering process is simulated by sampling the Spectre transient output at a high frequency that can represent a realistic digital counterpart of the original supply current, and then by filtering in the digital domain. Many filters with different properties have been examined, and it is concluded that the filter type has a minor

effect on the final accuracy of the predictions. A 4^{th} order type-II Chebyshev filter is used to generate the results presented here. The frequency response of the filter is shown in Figure 29. The filtered waveform is sampled at four times the IF frequency and the results are converted to 6-bit accuracy.

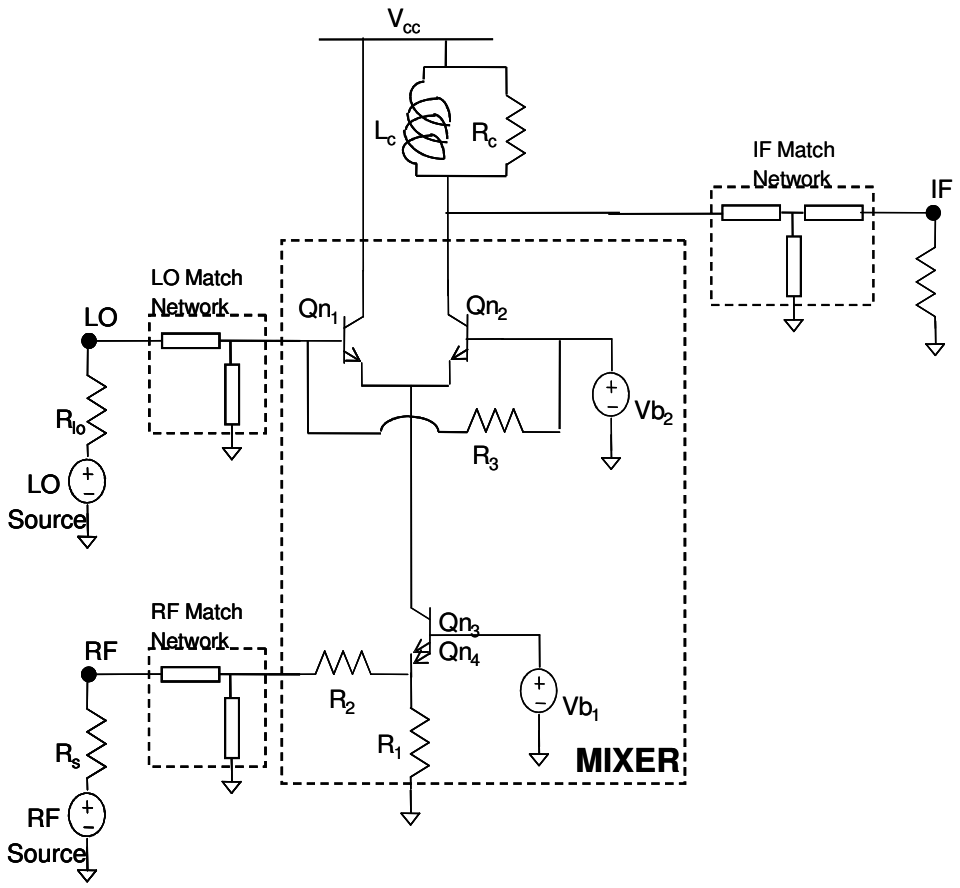


Figure 28: Bipolar RF downconversion mixer.

The alternate test stimulus is optimized separately for current and voltage signatures. In both cases, the search space for the input frequency is limited to a range from 820MHz to 1.2GHz. The LO is adjusted accordingly such that the IF is always 80MHz. In the first step of the experiment, the maximum stimulus is limited to 0dB, a high amplitude. Both voltage and current signatures settle for a stimulus with -5dBm amplitude at 918MHz. In the second step, the amplitude is limited to

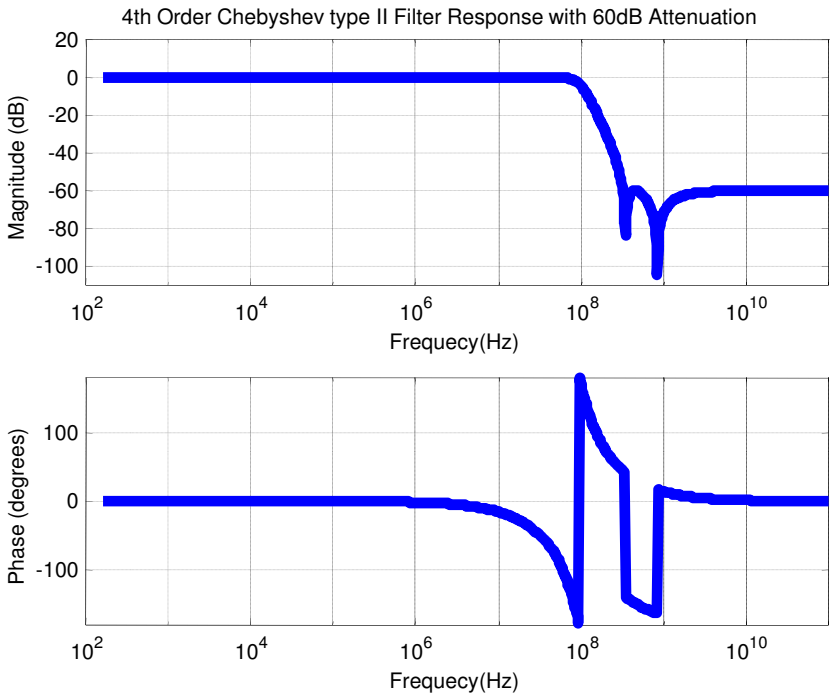


Figure 29: Frequency response of the digital lowpass filter.

−30dBm, which simulates a power restricted BIT situation. Voltage and current signatures settle for a stimulus with −30dBm but at frequencies 920MHz and 918MHz respectively. Table 3.2.4 shows the prediction accuracies for four specifications: IIP3, 1dBComp, CG at 920MHz, and PSRR at 920MHz. The first two columns list voltage and current based prediction errors for the input power of −5dBm. The maximum error for voltage-based prediction is below 0.5% while that of current is below 1%. The third and fourth columns list the same predictions with the input power of −30dBm. In this case, voltage based prediction error is more than 2.5%, while the current based prediction stays almost constant below 1%. Note that the 2.5% error comes from the IIP3, which is a non-linearity measure of the mixer. In this case, the input power is not large enough to make a significant difference in the voltage output, but the current signatures are still able to reflect strong features correlated to the specifications. Figure 30 shows the predicted versus actual specification values for the current-based test with −30dBm input. These results show that BIT of the mixer is feasible with a

single-tone alternate stimulus, and only current signatures achieve the same level of prediction accuracy at the lower stimulus power.

Table 5: Maximum percentage error in prediction: comparison for current vs voltage at −5dBm and −30dBm.

Type	Input (dBm)	IIP3 (%)	1dbCp (%)	CG (%)	PSRR (%)
Voltage	−5	0.448	0.196	0.042	0.017
Current	−5	0.807	0.218	0.324	0.104
Voltage	−30	2.660	0.258	0.078	0.035
Current	−30	0.885	0.230	0.343	0.128

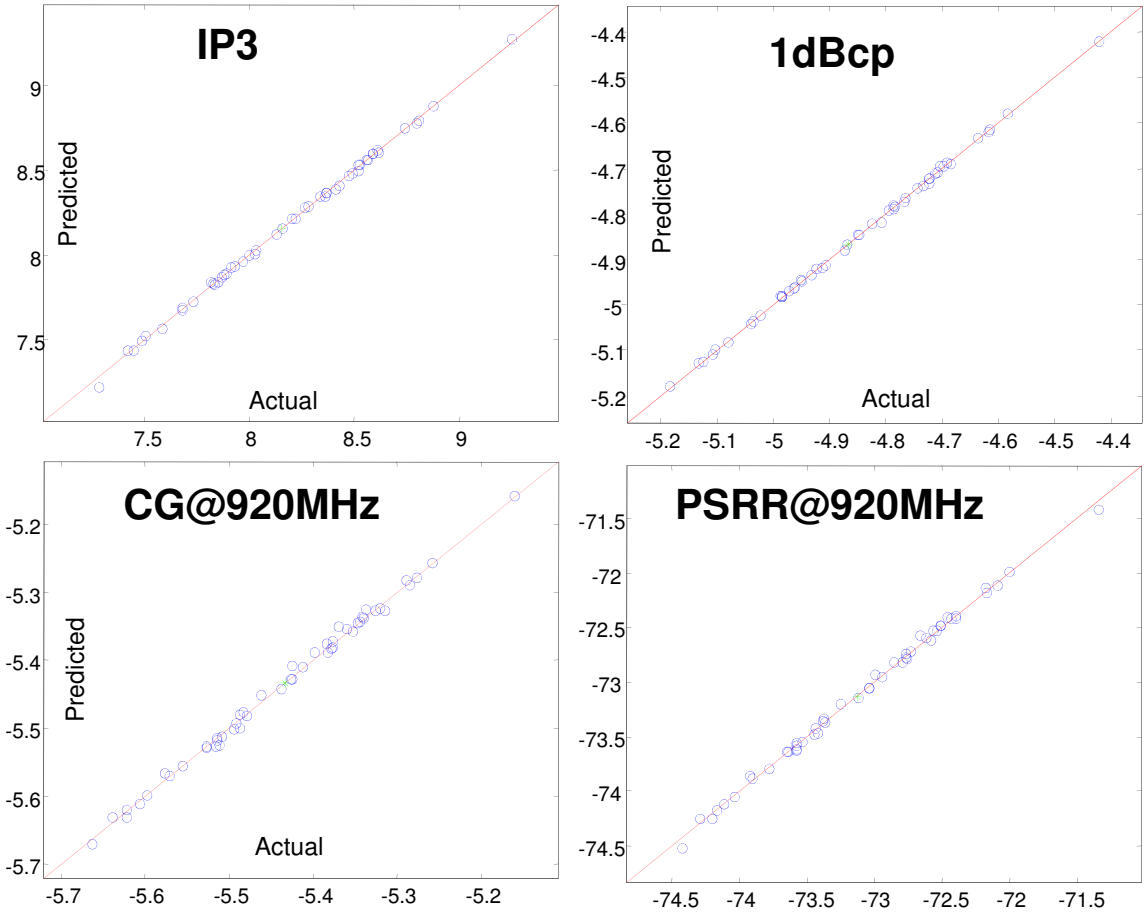


Figure 30: Actual vs prediction values for specifications of the bipolar mixer, red line shows the ideal mapping.

The second example demonstrates the non-intrusive nature of the current signature method. The scheme in [75] is employed as in Figure 31, a small resistor is added between the VDD and the DUT. Instead of a complex current sensor, we use a simple common source amplifier to perform I-to-V conversion and leave the rest to the alternate mapping module. This example is selected as a double-balanced downconverter mixer as opposed to a single-balanced configuration so that it will present an additional challenge. In a double-balance Gilbert mixer configuration, the currents through complementary branches, IF+ and IF-, add up to a constant. Hence, a single sense resistor and the associated current sensor can only relay information about the mismatches between the branches. However, most process variations affect a small region in the same manner changing all transistor parameters in the same direction. A single current reading cannot code much information about the parametric variations in this circuit. Instead, we implement a differential current signature by adding the same sense resistor and I-to-V converter to each complementary branch.

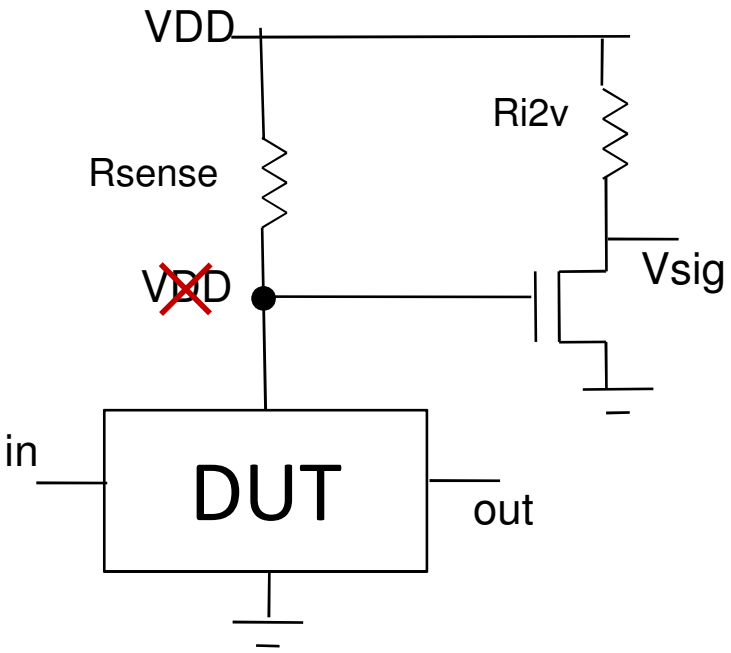


Figure 31: Sense resistor and I-to-V converter.

Figure 32 shows the 4GHz mixer designed for a 0.18μ CMOS technology with the proposed current sensor. Also note that resistors *Rsense* and *Ri2v* are preferred over PMOS loads because the original Gilbert mixer is only composed of NMOS transistors. If PMOS transistors are introduced, the sensor will no longer mimic the structures available in the DUT, but instead will show independent process variations imposed on PMOS variables. The goal is to keep the sensor structures as similar to DUT as possible. In this sense, each sensor has to be co-designed with the DUT. The specifications of the mixer are measured with and without the current sensor at a 3.97GHz input with a 3.96GHz LO. The corresponding CG, 1dBComp, IIP3 and NF are listed in Table 3.2.4. The effects of the proposed current sensor are negligible.

Table 6: Specifications of the mixer with and w/o current sensor.

	CG (dB)	IIP3 (dBm)	1dbComp (dBm)	NF (dB)
w/o Sensor	15.5	2.42	-6.89	13.6
with Sensor	15.2	2.23	-7.03	13.9

The 400-instance MC set is generated by 20% variation in *Tox*, *Xj*, *Vbx*, *Nch*, *Nsub*, *Xt*, *Rsheet*, and *Ls* as well 5% mismatch in *width*, *length*, *resistance*, and *inductance* values. Same variations and mismatch also applies to the components in the current sensor.

The following constraints are imposed on the GA-based alternate test stimulus optimizer: maximum two tones, total input power not to exceed -30dBm, input frequencies in the range 3GHz to 5GHz with matching LO, so that IF is centered at 10MHz, the current sensor load is limited to 50MHz bandwidth and the ADC has 10 bits of accuracy.

Only during the optimization progress, we realized a problem with the original double-balanced mixer design. The design was obtained from a third-party and was not properly verified with corner analysis. Further simulations revealed that for small

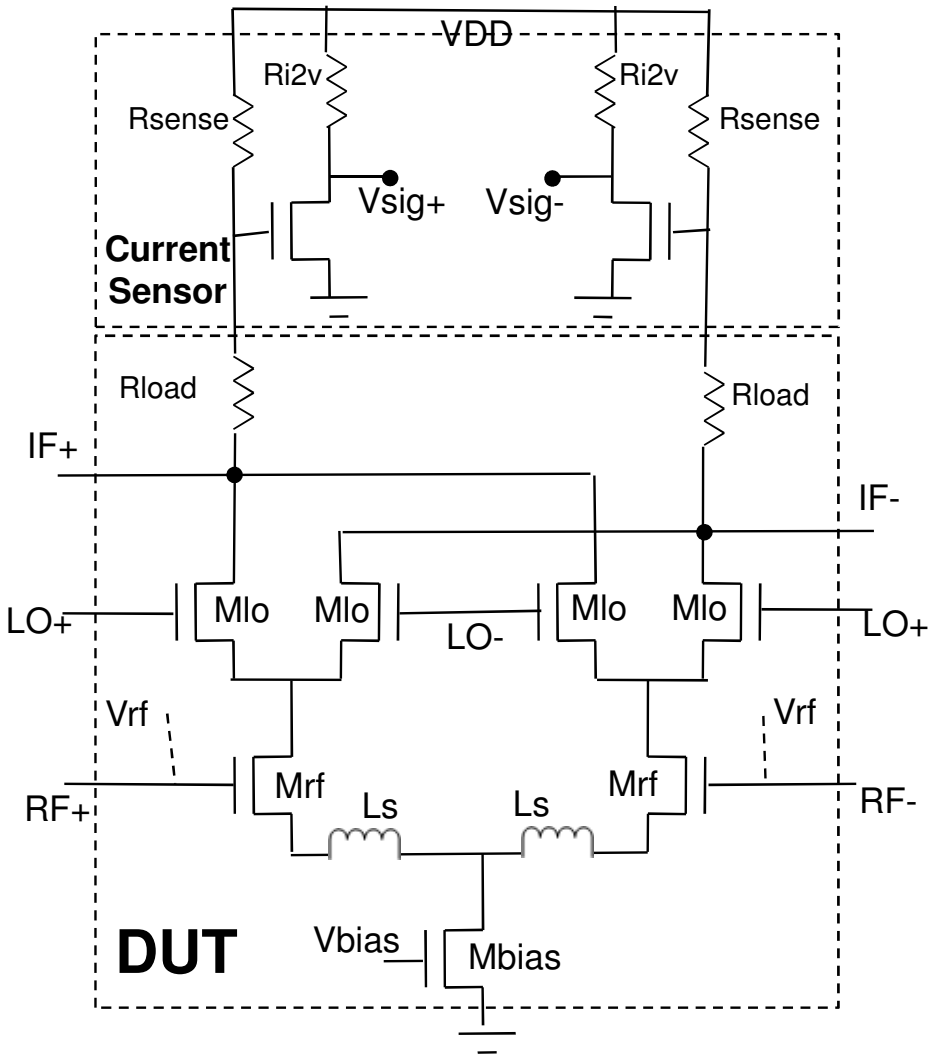


Figure 32: 4GHz mixer with sense resistor and I-to-V converter.

Ls variations, the mixer was unstable and showing oscillatory behavior at both IF output and the output of current detector. Instead of fixing the design, we took the challenge and checked if alternate test can flag these problematic instances. Out of 400 instances 102 were problematic. 44 of these instances oscillate with higher amplitudes and hence drive the ADC onto the rails leading to an easy detection scheme. Another 39 were identified by their additional large frequency components after the FFT. However, in order to eliminate aliasing effects, the sampling window had to be

increased considerably, which in turn increased testing time by two orders of magnitude. There are still 19 problematic samples which cannot be identified with these schemes. One possible solution is to study their oscillations with variable sampling windows, similar to wavelet analysis as in [45], we did not investigate this solution for the sake of simplicity. *The reader should keep in mind that these problematic instances were a result of faulty design process, a well-designed mixer does not show this kind of behavior, and otherwise the yield of the product would be extremely poor.* For the sake of demonstration, these instances are automatically discarded in Figure 33, which shows the predicted versus actual specification values for the remaining 298 instances. The maximum error is below 0.1dB, and the results show that the proposed current sensor can keep up with process variations as well as mismatch.

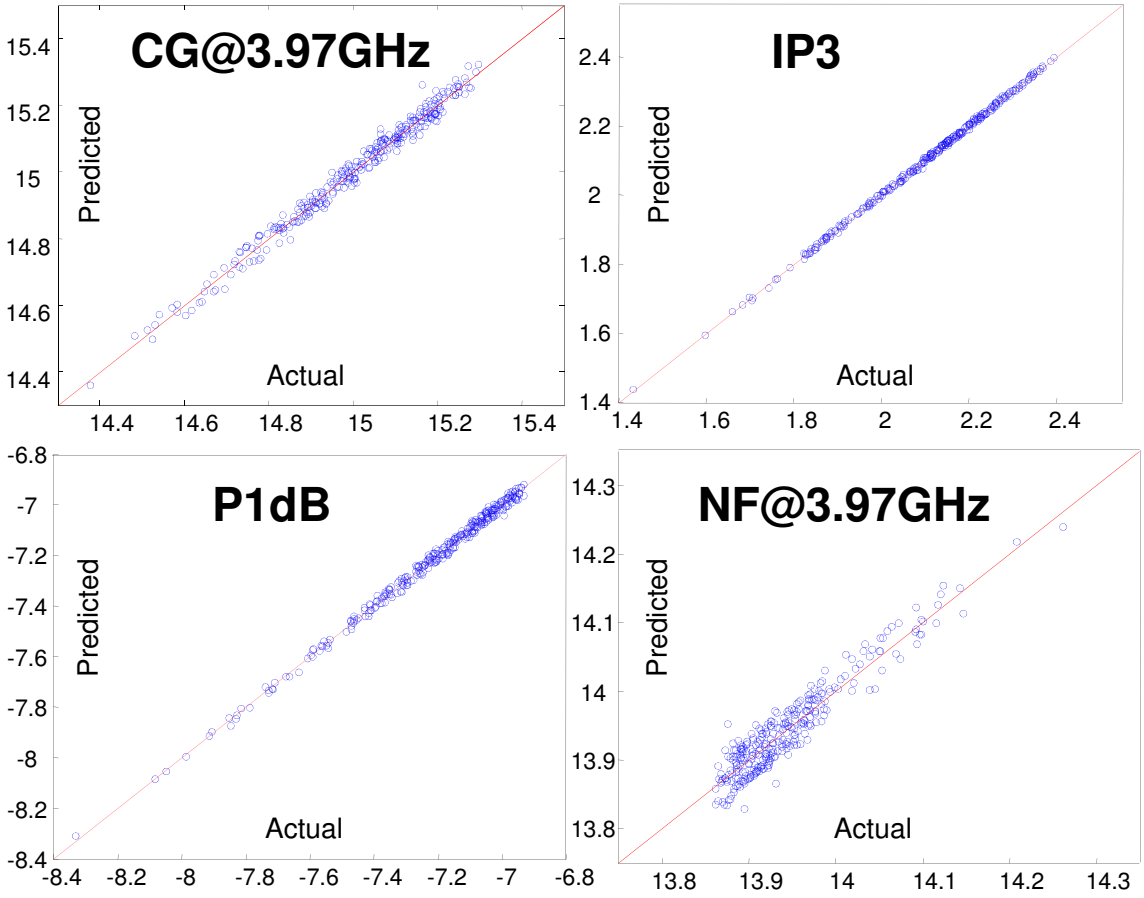


Figure 33: Actual vs prediction values for specifications of the mixer.

The two examples show that alternate test of RF components by current signatures is a non-invasive alternative to voltage-based BIT. The downside is that their application is exceptionally DUT specific and despite its simplicity the current sensor still needs to be co-designed with the DUT.

3.3 Alternate Test Generation with Behavioral Models

In this section, the main constraint in focus is the presence of intellectual property (IP) constraints. Built-in and sensor-based extensions of alternate tests make use of DfT features, which is only possible when the test engineer can influence the design process. Similarly, most other applications require access to a detailed device netlist and statistical information about the variations of process variables. However, such information is not always available, especially when the DUT includes third-party IP blocks. Similarly, the simulation models for complicated RF systems are generally very complex and time consuming for the repeated Monte Carlo simulations necessary during the alternate test stimulus generation process. Although the use of high-level behavioral device models has been proposed in the past to improve simulation speed, the complexity of these models limits the search space [76] or forces a computationally cheaper yet sub optimal greedy search [77].

The methodology in this chapter demonstrates a new variation of the alternate test methodology that makes no use of device netlists or process parameter distributions. A "gray box" approach is presented suitable for devices with IP blocks by creating a high-level simulation model from datasheet information and simple hardware measurements. This model is used together with a customized behavioral simulator to enable the efficient search of an alternate test stimulus that is optimal in terms of tester constraints, test time, and specification prediction accuracy.

The main contributions of this methodology are as follows: (i) new frequency domain simulator for test generation, which uses a simple model created with datasheet

information and limited hardware measurements to simulate the responses of a sample set of devices with different specification values; the models are only detailed enough to have accurate results for the alternate stimulus search; hence, they do not impose limitations on the initial search domain or force greedy algorithms; (ii) flexible ATPG structure for which the individual constraints of an ATE can be incorporated into the optimization process as parameters; (iii) GA-based ATPG method that codes the frequency content of the candidate stimuli in efficient gene formations; (iv) comparison of the alternate test and standard specification test in terms of accuracy, testing time, and equipment cost both on a benchtop setup and on ATE.

3.3.1 Traditional Alternate Test Generation Flow

Traditionally alternate test techniques have focused on test generation in the time domain with transistor-level device models [42–44, 46, 47, 78, 79]. Figure 34 shows this typical flow: in Part I, DUT netlists, semiconductor device models, and related process parameters are considered together to create a sample set of DUT models. Monte Carlo techniques are used to generate N instances which reflect the changes under given process parameter variation statistics. Then, these DUT models are simulated with conventional time-domain simulators, and their individual specification values are recorded. These actual specification values will be used partially as a training set for the supervised learner and partially as a validation set for comparison to alternate test results. In Part II, the alternate test stimulus, signature coding and mapping function are determined by a search loop. Depending on the nature of the search algorithm, the initial starting point may have an important impact on the final solution set. Using the candidate stimulus, the corresponding signatures are obtained through a conventional simulator for each member of the sample set of DUT models. Then the response signatures and the actual specification values for the training set are fed into a supervised learning algorithm which creates the final mapping function.

When the response signatures of the validation set are input to this mapping function, the result is a full set of predicted specification values for the validation set. These predicted specifications are compared with the actual specifications obtained in Part I, and the result gives a figure of test accuracy for the candidate solution set. If this accuracy satisfies the termination criteria, then the final test set is found; else, the search loop runs again with the new stimulus. The computational complexity of the methodology is dominated by the loop in Part II: $O(Q)O(Nn_l)$ where N is the size of the sample set, n_l is the total number of candidate stimuli considered throughout the search loop, and $O(Q)$ represents the computational complexity of one simulation. For a conventional simulator, it can be assumed that the total simulation time dominates over the time spent for generating the mapping functions.

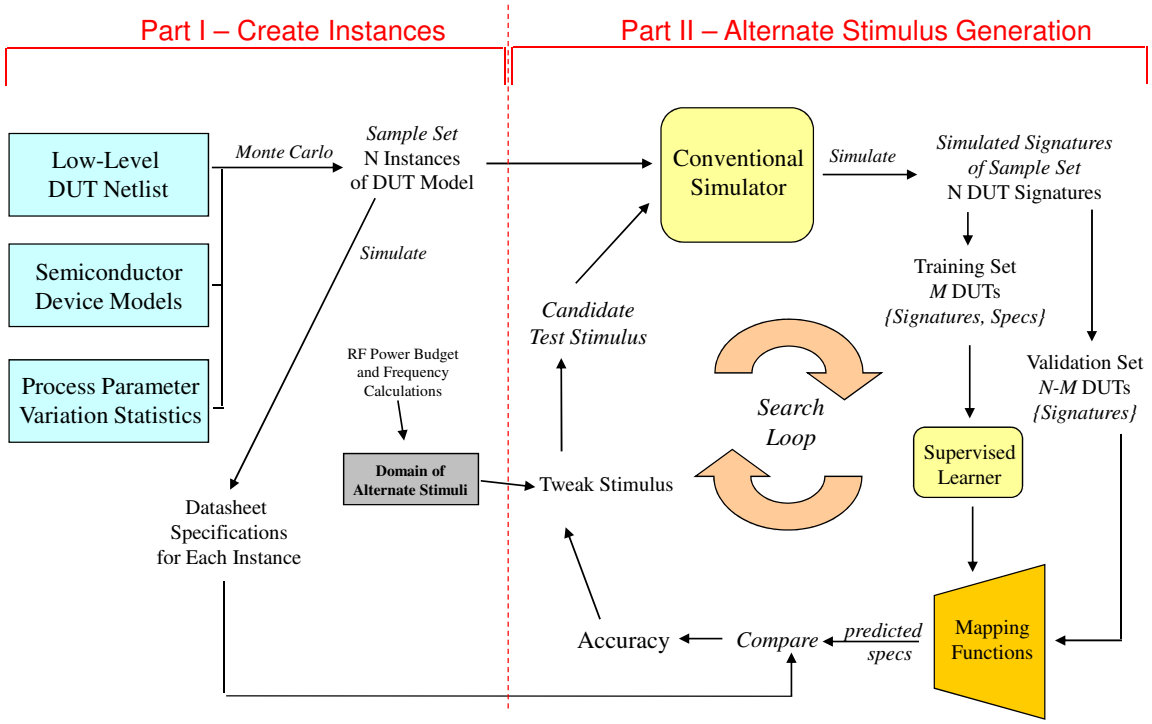


Figure 34: Crafting an alternate test stimulus; Part I: create sample set instances and measure reference/actual specification values, Part II: alternate stimulus generation with conventional simulator.

3.3.2 Alternate Test Generation Flow with Behavioral Models

Considering that simulations will be run on a statistically sound set of devices, the computational complexity of the search algorithm becomes prohibitive. This is mainly due to $O(Q)$; the circuit simulation time is cubically proportional to the number of nodes and to the number of voltage and current variables [80]. By moving from transistor level netlists to higher level behavioral models, the simulation time for a complex RF device is reduced roughly by two orders of magnitude [81]. Replacing transistor-level models with behavioral models not only reduces the complexity of the algorithm, it is also the only solution when IP blocks are present in the DUT. The netlist and statistics of process parameter variations are not available, hence conventional simulators and sensitivity analysis [78] cannot be used.

In this work, a high-level behavioral model of the RF system is created by a gray-box approach, where a set of N behavioral level instances are created from hardware measurements on N devices. Consequently, stages in Part I of Figure 34 are replaced with data from hardware based measurement. With no models to apply MC analysis, the specifications of N devices are measured one by one by using classical test equipment. When compared to the original alternate test flow, these hardware measurements do not present an overhead in terms of test development time or cost; because the exact same measurements are already required for hardware calibration of the mapping functions. N devices are divided into training and validation sets such that each set is close to being a representative sample of the overall specification distribution.

Since the computational complexity heavily depends on the number of instances generated, $O(Q)$ needs to be kept at a minimum without compromising the quality of the generated test. Therefore the alternate test stimulus and the response signature are based on a carefully selected set of test response features. Frequency domain features prove more efficient in this case, where only the propagation of a number of

tones needs to be simulated. This behavioral flow is shown in Figure 35.

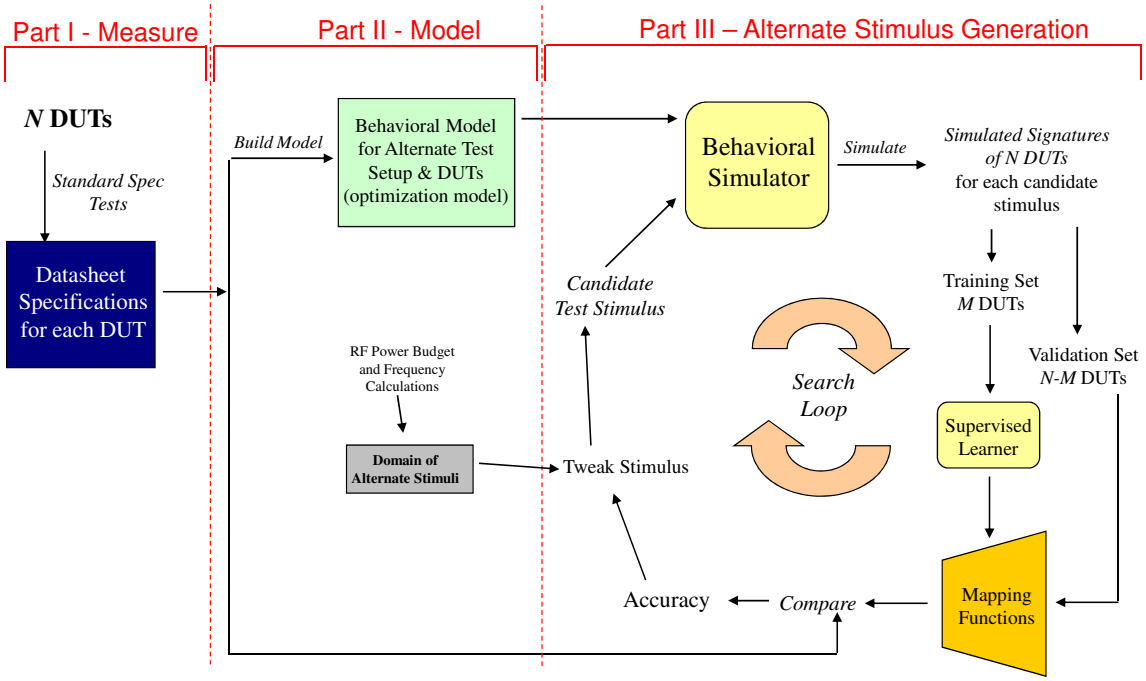


Figure 35: Alternate test stimulus generation with hardware measurements and behavioral models. Part I: datasheet specifications are measured for a sample set, Part II: a behavioral model is created, Part III: alternate stimulus generation with the behavioral simulator.

3.3.3 Generating and Evaluating Mapping Functions

The performance metrics for each DUT are obtained by mapping the signatures into the specification space. These mapping functions are constructed by a supervised learner. In the literature, MARS has been the typical learner used with alternate tests. The final functions can be visualized as a weighted sum of basis functions made of splines. MARS uses an initial recursive partitioning during training to gradually add these basis functions using forward stepwise placement; then, a backward procedure is applied and the basis functions associated with the smallest increase in the least squares fit are removed [48]. A detailed description of MARS applications in alternate test is provided in [39].

In this work, three other supervised learners are used and the results are compared in following sections: least angle regression (LARS) is a computationally efficient forward selection method as described in [82], the experiments in this work use R packages available from the same authors. Multiple additive regression trees (MART) is an implementation of the gradient tree boosting methods for predictive data mining, R packages are available from [83]. Sequential minimal optimization (SMO) is a training algorithm for support vector machines as described in [84], software available from [85].

The quality of a mapping function can be judged by many ways, for example by checking the correlation coefficient *adjusted* R^2 (R^2_{adj}) between the actual and predicted specification values of a validation set. With any supervised learning tool, the final mapping functions will only be as good as the quality of the data used for the training set. In this case, the two components of the training data are (i) the actual specification values coming from hardware measurements, and (ii) the response signatures from the simulator. The prediction error in alternate test err_P is dominated by the repeatability error of the actual specification measurements rep_A. In this section, we use the difference

$$AR = -max(err_P - rep_A), \tag{6}$$

as a figure of merit for alternate test quality. The maximum is computed over all specifications and all DUTs in the validation set. It is important to understand that a large value of AR may not present a problem if the performance metrics for the DUT are far away from the specification pass/fail boundary. In that case, DUT may as well be classified correctly even if predicted performance metrics do not closely match the actual ones. R^2_{adj} does not help either in this case, as it is a measure of average err_P. A more robust quality check is based on pass/fail, or binning, classification accuracy, hence accurately accounting for yield loss and test escapes. The following formula

weighs test escape heavily over yield loss and can be used for judging test quality:

$$CR = min(alt_{TestEscapes} - act_{TestEscapes}, 0) - \frac{min(alt_{YieldLoss} - act_{YieldLoss}, 0)}{1000}, \quad (7)$$

where *alt* represents misclassifications with the mapping function, and *act* represents misclassifications with actual hardware specification measurement. However, driving a search algorithm based on a discontinuous variable like CR has poor convergence implications. In following sections, we also compare the efficiency and success of alternate ATG driven by AR versus CR.

In this work, for reasons of speed, a custom behavioral simulator is developed in Matlab, which is based on the propagation of frequency domain tone information through a series of elements representing the devices, transmission lines, equipment interfaces, and cables in the actual test system. It can model typical RF module behavior, including amplification, generation of intermodulation products, frequency translation, compression, filtering, harmonic distortion, and feed-through. Phase noise from the LO is also represented by a bunch of spurs around the fundamental LO tone.

The simulator only uses magnitude information of the frequency spectrum. The phase is neglected in a controlled manner, because by doing so: (i) faster simulation times are possible by reducing the information to be processed at each step; (ii) the modeling procedure is simplified, and models can be generated by using a spectrum analyzer and a multi tone signal generator; (iii) the inaccuracies resulting from addition of two tones with different phases is compensated by not considering the resultant tone in the final signature, this generates a robust alternate signature; (iv) it can make use of free running local oscillators for up/down conversion.

The proposed simulator is fast enough so that the simulation time for a total of N instances is as short as the time it takes MARS to create the corresponding mapping functions. Hence, the complexity of the methodology is given by $O(Q'+M)O(Nn_gn_p)$, where $O(Q'+M)$ is the total complexity of the proposed simulator and the creation

of mapping functions. The accuracy and speed of the simulator, frequency domain alternate response "testimator" (FDART), is compared with two general purpose off-the-shelf simulators: time domain (TD) example is the Cadence SpectreRF with periodic steady state (PSS), and frequency domain example is the Cadence SpectreRF with harmonic balance (HB). These results are presented in the next section.

The test path is divided into a series of elements for modeling purposes. Any element with a lookup table is interpolated linearly for missing values and extrapolated with the exact same end-value closest to it. The elements are determined as follows:

- loss is represented by a filter element made up of a lookup table for frequency loss value pairs;
- any input (LO and other RF signal sources including stimuli) is represented by an input element, a group of tones corresponding to the harmonics and their phase noise tails;
- frequency translators are represented by a mixer element made up of a normalized matrix and correction tables. The normalized matrix represents the output corresponding to a single unit amplitude tone as a result of a predefined number of (defined by the model) LO and input harmonics varied by input frequency. Hence, the normalized matrix generates LO±IN, LO±2IN, LO±3IN, ... 2LO±IN, 2LO±2IN, ..., LO feedthrough (LO, 2LO, ...), and input feedthrough (IN, 2IN, ...). Once these tones are generated, three correction lookup tables scale the output with respect to LO frequency, LO power, and input power;
- amplifiers are represented by amplifier elements made up of three lookup tables: frequency-gain value pairs, gain compression, and intermodulation (IMD). Intermodulation products are considered up to an order defined by the model.

The frequency domain is quantized in terms of both frequency and amplitude, the minimum quantization step is determined by a simulation parameter. If two tones with the same frequency are generated by an amplifier element or a mixer element, then these tones superpose and their powers add up as if they are in phase. If one tone is significantly larger than the other, which is evaluated by a limit defined by a simulation parameter, the error coming from the phase difference will not be significant. If the two tones are comparable in magnitude, then the resultant sum is marked by a dirty bit. In this case, the dirty bit propagates into latter elements and marks related tones and IMD products accordingly. At the end of the chain, only the tones with clean bits are considered for the alternate test signature.

Cables, transmission lines, socket interfaces, and all kinds of filters are directly converted into filter elements. A passive mixer is converted into a mixer element followed by a filter element. An active mixer is a mixer element followed by an amplifier element. The modeling process is simply composed of dividing the test path into a logical chain of these elements, and deriving the corresponding matrices and lookup tables using a spectrum analyzer and up to two signal generators, one for input and one for LO, where applicable. Then, each component is defined by a software structure in Matlab containing the appropriate tables and matrices. The elements are connected in the form of a graph notation. The coding of the simulator is performed such that it makes use of the efficient matrix operations in Matlab by combining many instances of the model into a single matrix and computing responses simultaneously.

3.3.4 Search Algorithm

The search algorithm drives the optimization loop in Figure 35 and directly impacts the efficiency of the ATG. Although greedy algorithms can cut down the computational complexity by reducing n_l, that is by considering a single candidate stimulus

per execution of the optimization loop, their application is limited to certain domains [86]. Preliminary studies and experience shows that the RF stimulus domain presents a neither monotonous nor well-behaved topology, the former because of saturation and intermodulation effects, and the latter resulting from isolated frequency bands from filter effects. As a result, greedy algorithms easily get stuck in local minima. Considering the size of the problem space, brute force approaches are also not feasible. On the other hand, GA have been shown to be very powerful for similar applications [87]. In the next section, we compare the efficiency of three algorithms. The first one implements a random walk (RndW). A number of, K_s, initial candidate stimuli are generated by randomly assigning number of tones, tone frequencies, and tone amplitudes. In each iteration, the frequency and amplitude values are perturbed slightly but randomly. After a fixed number of iterations, N_s, the candidate with best quality is selected over all considered in any step of the process. This random walk implementation provides a baseline for the performances of other more structured search algorithms. The second implementation is a gradient-based search (GradS) as described in [28]. A number of, K_s, initial candidate stimuli are generated by randomly assigning number of tones, tone frequencies, and tone amplitudes. In each iteration, the frequency and amplitude values are changed in the direction of the greatest descent in cost function. When the algorithm converges, that is a local minimum is reached, the number of iterations is noted. In order to make a fair comparison between this algorithm and the others, the starting point is reseeded and search restarted until a total of $N_s \times K_s$ iterations are completed over all groups. The candidate with best quality is selected over the best of K_s groups, and the number of iterations within that group is recorded as a measure of efficiency. The third implementation is a GA-based search (GAbS) as described below.

GA emulates the natural selection process: a large number of individuals form a population; each individual is evaluated for its fitness, and the ones with higher fitness

values have a higher probability of bearing children; the less fit individuals are replaced with these children so survival of the fittest strains are guaranteed; on the same while a small amount of mutation rate creates children with new fitness conditions; the population evolves slowly but in a steady pace. GA-based search algorithms cannot guarantee global optimum in finite time, yet they tend to easily move away from local minima. The downside is that GA is computationally more demanding, because n_l becomes a product of n_p, the number of individuals in a population, and n_g, the number of generations required to satisfy optimization constraints. In our experiments, $n_p = K_s$ and $n_g = N_s$, to make a fair comparison of search algorithms.

The quality of the GA implementation depends mainly on two factors: (i) a concise gene notation which transforms the information each individual possesses into a code for exchanging that information without breaking its meaning; (ii) a fitness function which can identify the better individuals from the lesser without losing diversity by over penalizing others. For this study, we have experimented with a number of fitness functions which are based on different definitions of test accuracy. Although different functions yielded similar optimized stimuli at the end, the one based on rms error with 0.05 significance interval required less number of generations to produce the same level of quality. The gene notation is also customized to this application in the sense that it speeds up the convergence process. Since the candidate stimuli are essentially multi-tone signals, each individual is represented by a D-bit gene sequence. Each b consecutive bits, called as gene bytes, represents a quantized amplitude value for equally spaced frequency components with spacing Δf_{min}. The simulator interprets any amplitude below A_{border} as a no-tone, which is an infinitely small tone not represented in an input element. The whole input frequency domain is divided into K overlapping regions, limited by the bandwidth available on the signal generator and defined by the center frequency component.

Figure 36 shows the mechanisms for generating new children. In crossover (denoted with a), a random gene byte location is selected (red dotted line); the first parent's bits from the right of the selected location are combined with the second parent's bits from the left of the location. Each region designated by the center frequency component can only mate parents within; this way the newborn (new child a) does not violate the bandwidth limit. The other main mechanism, mutation, makes a random amplitude difference in one of the quantized frequency locations in such a way that a tone already smaller than A_{border} becomes larger (new child b) or vice versa (new child c). The GA implementation also lets a small percentage of *elites*, individuals with best fit scores, propagate to the next generation without any modification. The use of the elites guarantees that best fitness value will not get worse from one generation to the next.

We also implemented a migration feature in the GA. Every n_{mg} generations, copies of evolved individuals from different regions are subjected to circular migration by changing their center frequency with that of the region they are being migrated to. This way, a frequency-amplitude pattern that thrives in one region has a chance to mate with another leading pattern in another region. The reader can refer to [87] for the theory behind these features.

3.4 Simulation Experiments

In this part of the study, several design factors for the test generation algorithm are investigated with two simulation examples. The first example is a 900MHz receiver front-end; LNA and mixer designs are as shown in Sections 3.1.4 and 3.1.5. The specifications of interest are gain, IIP3, 1dBComp, and NF. The second example is a transmitter front-end; quadrature modulator, voltage controlled oscillator (VCO), and RF amplifier designs follow [88], the power amplifier and the differential-to-single-ended converter (D2S) are based on [89]. The specifications of

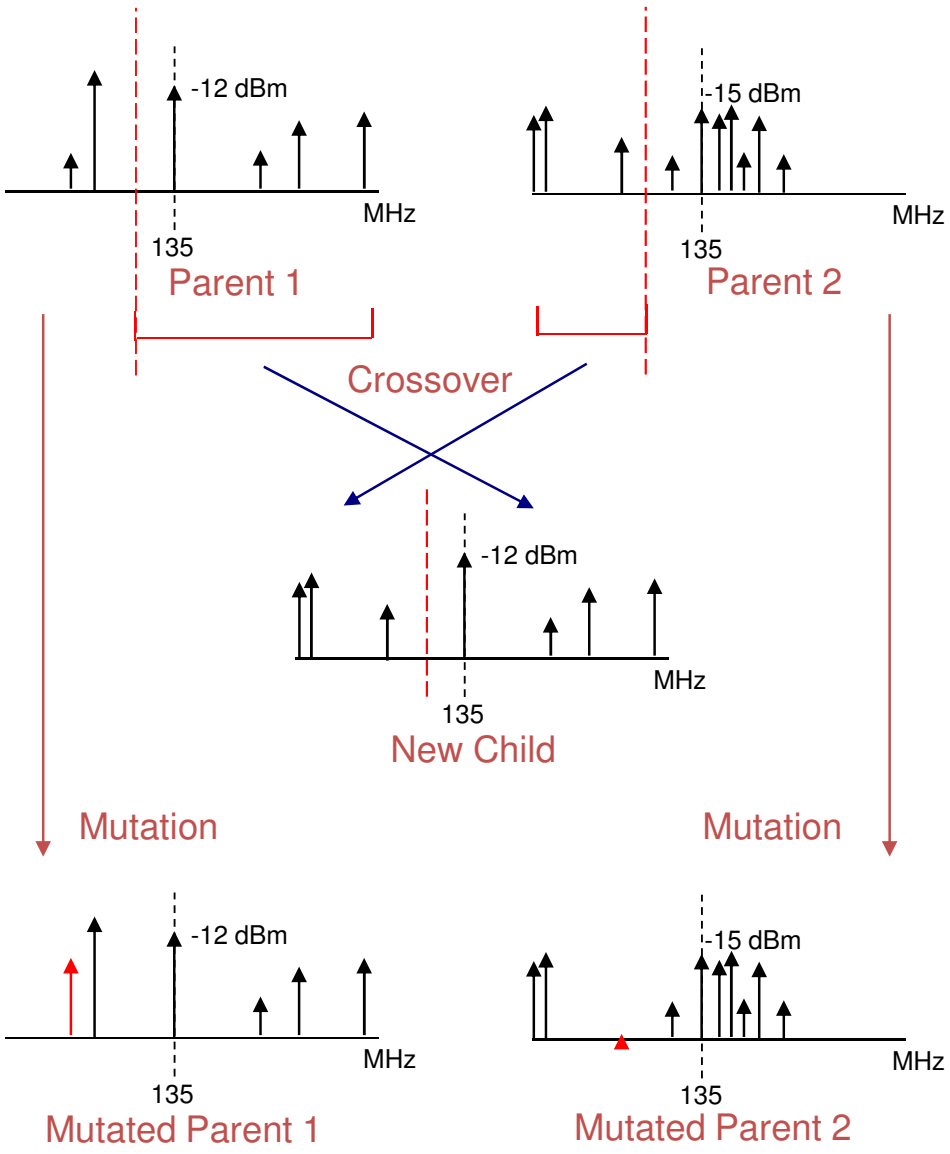


Figure 36: Crossover (a) and mutation (b, c) in GA with the proposed gene notation.

interest are gain, output-referred third order intercept point (OIP3), 1dBComp, and LO spur.

Three simulation setups are used to capture specifications and emulate test conditions. First, conventional simulation methods are used for determining specifications of the set of devices built with full-blown transistor-level models and process variations. This mode emulates conventional test methods as shown in Figure 37a for

the receiver and Figure 37b for the transmitter. Second, same transistor-level models are used with an alternate test fixture as shown in Figure 37c and Figure 37d. The up-conversion mixer before the receiver enables a low-frequency multi-tone input source, and the low-pass filter (LPF) limits output sideband for sampling. Similarly, by using a down-conversion mixer and LPF at the output of the transmitter, a cheaper IF sampler condition is emulated. Third, DUTs and test fixtures are replaced with behavioral models. In this mode, simulations are performed in the frequency domain. Quadrature channels of the transmitter are handled as two independent paths, replicating all components after the demodulation mixers.

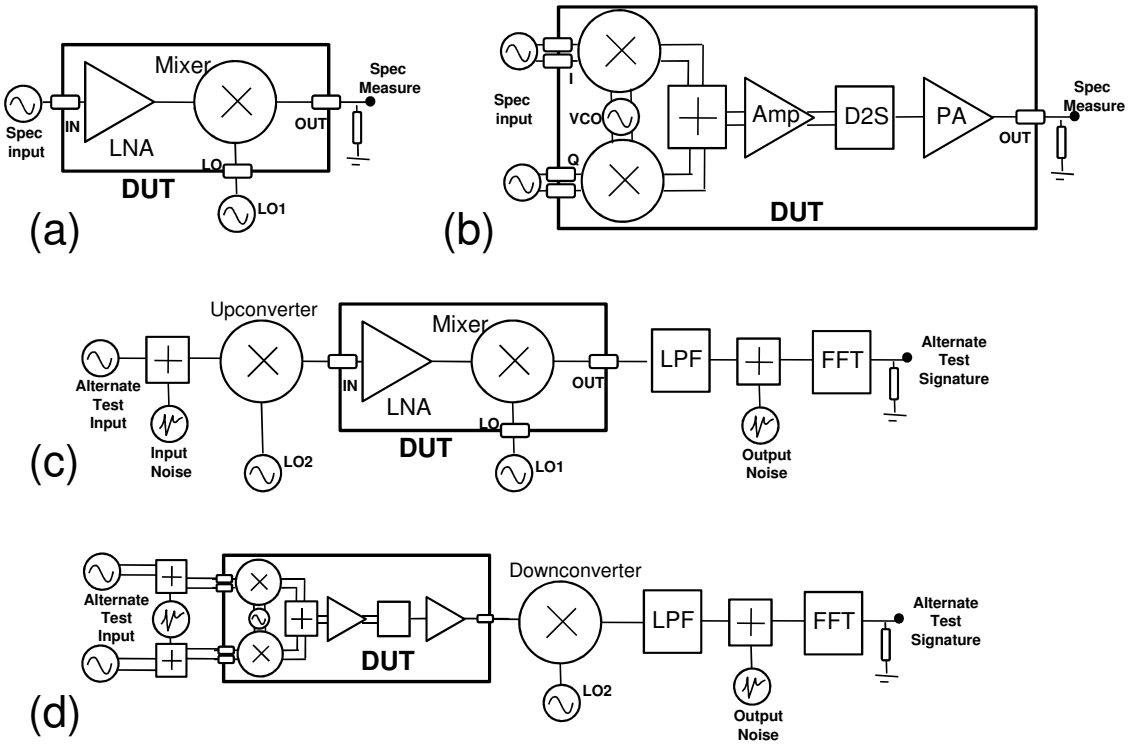


Figure 37: Simulation setups. In part (a) receiver front-end is composed of a LNA and mixer, in part (b) transmitter front end is a quadrature modulator followed by a RF amplifier, a differential to single-ended converter and a power amplifier. Parts (c) and (d) show alternate test setups for transistor-level simulations.

3.4.1 Design Factors on Test Generation

Several factors impact the effectiveness of the test generation algorithm. The most important three are identified for this study.

- Sample size: Supervised learner depends on the availability of a training set that closely represent the entire DUT population. Although it is common sense that a larger set has a higher probability to sample a wider range of the population, it also increases simulation time. For a fixed amount of total simulation time, increasing sample size will result in less search iterations, and sub-optimal test patterns. In order to understand this trade-off, three sample sizes are evaluated: small (50 instances), medium (500), and large (5000).

- Distribution: Larger sample sizes do not show benefits unless the population is sampled in the proper way. The distribution functions governing simulation of process parameter variations is as important as sample size. Five different distribution functions are investigated for Monte Carlo analysis: all parameters linear (LIN), all parameters Gaussian (NORM), all parameters mixture of linear and Gaussian in a 1:3 sample ratio (1:3MIX), linear around response surface model (RSM) corners (RSMD), and Gaussian around RSM corners (NARD).

- Noise: An ATG algorithm needs to be robust under input and sampling noise. Since alternate ATG depends on a supervised learner, it can be over-trained on training examples and exhibit sensitivity to noise. Three levels of output noise is simulated by adding random numbers between $\pm 0\%$ (clean), $\pm 2\%$ (typical), and $\pm 10\%$ (noisy) of the actual value.

3.4.2 Experiment Setup and Results

The combination of design factors create a total of $3_{samplesizes} \times 5_{distributions} \times 3_{noiselevels}$ = 45 training sets. These training sets are then matched with combinations of supervised learners, AR versus CR, search algorithms, and simulation engines as was listed in the previous section for a total of $45 \times 4 \times 2 \times 3 \times 3 = 3240$ cases. In each case, the same validation set is used to record AR and CR. This validation set is created independently from training sets with $18,000$ samples, half distributed as *1:3MIX* the other half as *NARD*, and with equal partitioning for 0%, 2%, and 10% noise budgets. The repeatability is assumed to be perfect for the simulation examples; hence,

$$AR_{ValidationSet} = -max(err_P). \tag{8}$$

Also it is assumed that actual specification measurement tests have no test escapes or yield loss; hence,

$$CR_{ValidationSet} = -alt_{TestEscapes} - \frac{alt_{YieldLoss}}{1000}. \tag{9}$$

Obviously, it is not possible to cover all $3,240$ parameter interactions in this thesis, hence only a handful of interesting ones are examined. First, all cases are sorted by AR and CR, and top 100 is studied for first-order pareto items. The selected search algorithm tops the overall pareto. *GAbS* performs up to 64% better than *RandW* and *GradS* under identical conditions. There are few cases that *GradS* performs better, but when we consider the number of iterations for *GradS* to converge for that subgroup, it is more than 8× that of generations in *GAbS*, and none of those cases make the top 100. The other factor that stands apart is the amount of noise. Looking at top 100 cases, 98% of them are training sets with 10% noise. Further investigation is required to understand if the result would be the same when validation set maximum noise budget is limited to 2% instead of 10%.

Another interesting observation is on the choice of simulation engine versus simulation time required, and the accuracy of results. There is roughly 40%/40%/20%

distribution in the top 100 accordingly for *TD*/*HB*/*FDART*, which means time domain and harmonic balance engines are roughly equivalent in accuracy reflected to end results. What *FDART* lags in accuracy is made up in simulation time. For the 20% of the top 100 that uses *FDART*, simulation time is 11× to 134× shorter than *HB*, with an average of 88×, and 57× to 1602× shorter then *TD*. To evaluate the benefits of a faster engine, another experiment is performed where simulation time is limited to that of *TD*, instead of using constant number of iterations for each engine. Using *AR*-driven *GAbS* with *MARS* and 10% noise, *FDART* yields 6× better *AR* than *HB*, and 28× better than *TD*. Although the gain in simulation time does not linearly impact *AR*, it is a significant trade-off.

When it comes to choosing a fitness function, it is the combinations that matters. *AR* benefits more from *MARS*, while *CR* shows better results on *SMO*. This observation can be explained by the nature of the supervised learners. *SMO* is based on support vector machines and is a classification tool modified into a regression learner. *MARS*, on the other hand, is built on splines that are specifically chosen to yield continuity between pockets of best fitting regions. While *CR*-driven cases show improvement under a more advanced distribution scheme such as *NARD*, *AR*-driven cases are indifferent between *1:3MIX* and *NARD*. Also observed is that *1:3MIX* copes better with medium sample sizes than *NARD*, hence, when *AR* is the driver, the cut in simulation time can be better spent on a *HB* engine rather than *FDART*.

3.4.3 Key Summary on Alternate ATG

Looking at the trade-offs observed in the previous sub-section, the golden recipe for alternate ATG depends on the depth and amount of information available about DUT. Three example cases will be studied here.

- For a case with in-house fabrication capabilities, typically test engineers will have total access to transistor-level models and extensive information about

the process variations. Availability of strong feedback back to process control enables product engineers to stick with pass/fail or bin boundaries originally planned for market needs; hence, bin boundaries change little from design step to shipment of the product. Also, early engagement in fabrication process means more time for test generation, and ability to undertake more computationally demanding recipes. The golden recipe under these conditions is a GAbS driven by CR, simulated with a HB engine on a large sample size distributed with NARD, and a SMO-type supervised learner.

- A fabless design house with access to transistor-level models may not always have critical feedback to process control. Consequently, pass/fail or bin boundaries may as well be a moving target. Also, typically, less time and iterations are available for HVM test generation. For this case, golden recipe consists of GAbS driven by AR, simulated with a HB engine on a medium sample size distributed with 1:3MIX, and a MARS learner.

- For a case with third-party IP blocks, typically time-to-market is the key, hence time available for test generation is limited. In this case, FDART is the engine of choice, and it favors AR-driven GAbS on a large sample size with MARS learner. Sample distribution cannot be adequately controlled due to IP blocks, but this combinations performs equally well with 1:3MIX and NARD.

3.5 Hardware Experiment with IP Blocks

In order to validate the proposed approach, a receiver front-end chip from RF Micro Devices is selected. RF2411 [90] is a dual-band LNA followed by a mixer with pins for access to both components individually as shown in Figure 38. As we have no access to transistor-level models or process distributions, this is a challenging example for alternate ATG and a good case to showcase FDART. For the sake of simplicity, only one band (850MHz) is considered for thirteen specifications: LNA gain, input standing

wave ratio (ISWR), output standing wave ratio (OSWR), reverse isolation (RevIso), IIP3, NF; mixer gain, ISWR, IIP3, NF; and cascade (LNA + Mixer) gain, IIP3, and NF.

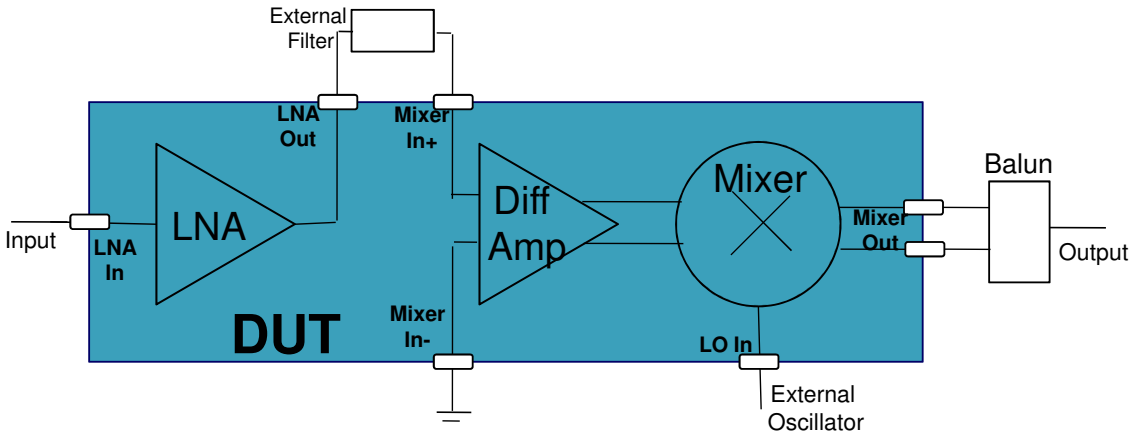


Figure 38: Receiver front-end chip as DUT. LNA and mixer specifications are handled stand-alone as well as cascade. Differential mixer is converted to single-ended input/output. External filter and balun are not considered as part of the DUT to simplify the experiment.

The sample size is selected based on the time limitation imposed by hand socketing the devices. More than five hundred ICs are bought in batches over a time interval of two months in an effort to partially sample lot-to-lot variation. However, all DUTs pass RFMD's high-volume manufacturing screening, hence no artificial pass-fail boundaries are introduced by the author, the results are studied by *AR*.

The experiments can be studied in three steps.

- Standard specification tests: these benchtop tests are performed to evaluate the specifications of each IC with the conventional methods and obtain equipment cost, testing time, and repeatability measures for the standard tests.

- Benchtop alternate tests: the standard specification test is replaced with an alternate counterpart and the results are compared.

- Alternate test on ATE: the benchtop alternate test is migrated onto a commercial ATE platform for a feasibility study.

To demonstrate the equipment cost reduction resulting from alternate testing, the target ATE is selected as a mixed-signal open architecture [91] tester instead of an RF tester. The tester can create 12-bit arbitrary waveforms (AWG) at 800MSps, yet limited to 200MHz bandwidth because of clock bleeding. It can also coherently sample at 800MSps with 12-bit accuracy.

3.5.1 Standard Specification Tests

Testing the listed thirteen specifications of a single IC requires seven test configurations involving different equipment and pin connections. To minimize the repeatability problem, we designed a loadboard that can automatically cycle through these configurations with a single IC insertion. The various test equipments connect to the loadboard and routed to the proper I/O of the chip through relays. Figure 39 shows this setup: the IC is interfaced through a socket on the socket interface board (SIB), which has five RF connectors for LNA input, LNA output, mixer input, mixer output, and LO input; the test equipment consists of a spectrum analyzer, a network analyzer, two signal generators, a noise source, and a power supply; all the equipment connect to the relay board (RB), and routed to the proper input and output ports on the SIB; the relay driver board (RDB) -figure shows the unpopulated stage- hosts relay controller chips, for which the digital control is interfaced to a laptop using a NIDAQ card; the whole process is automated by a custom written Agilent VEE program, which controls the test equipment through GPIB. Figure 40 shows a snapshot of the VEE code, and Figure 41 shows the user interface.

The partitioning of the loadboard into three parts provides major benefits: (i) the SIB can be reused in alternate test setups, hence standard and alternate test results

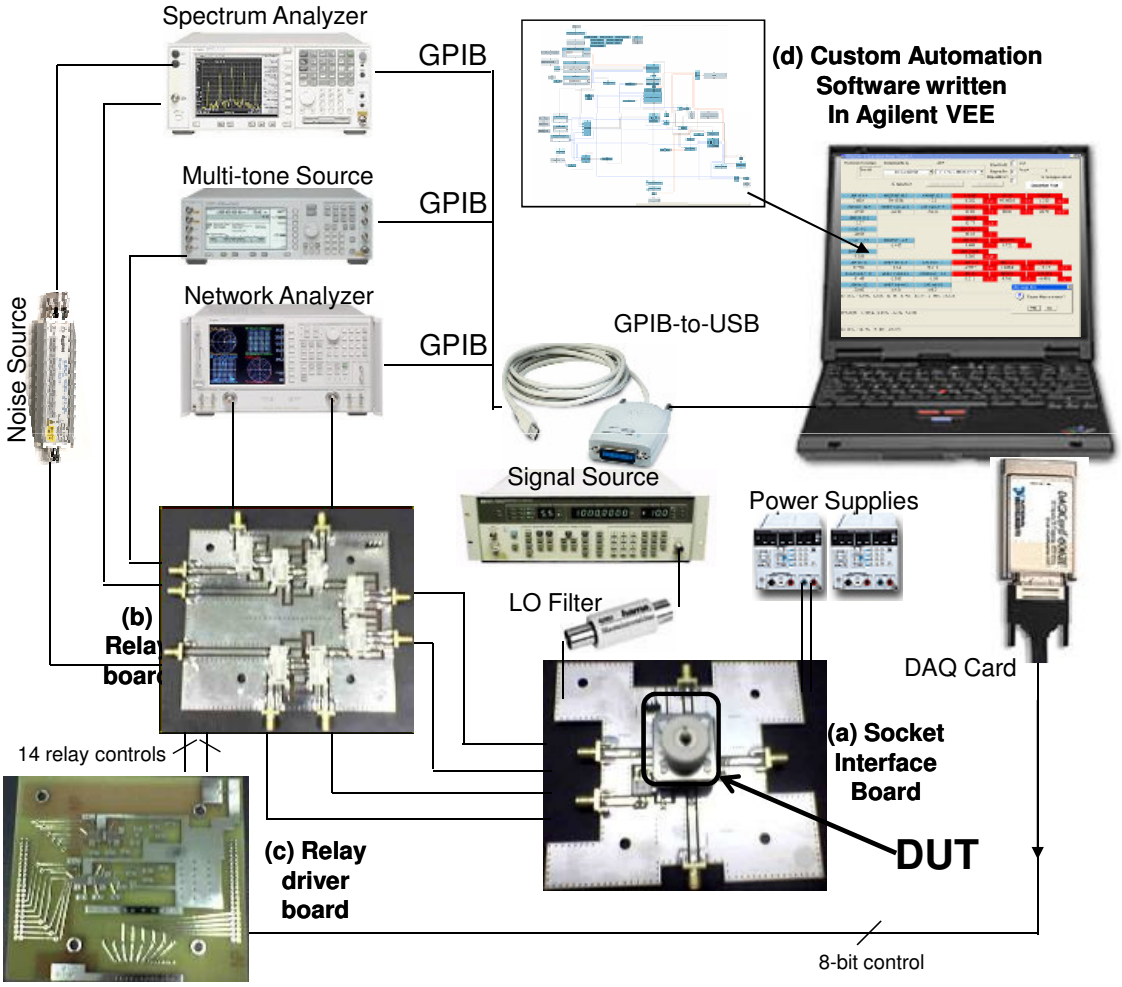


Figure 39: Standard specification test setup. DUT is placed in the hand socket shown on the (a) socket interface board. This board is connected to measurement equipment thru relays on (b) the relay board. (c) Relay driver board controls these relays. Synchronization and logging is performed by (d) custom automation software.

can be compared directly; (ii) the separate RDB isolates the sensitive RF paths on RB from the digital control. The three different boards are vertically connected with cables and encapsulated into a Faraday cage, which is necessary for NF measurements.

The accuracy of the proposed standard specification tests are also measured by a number of repeatability experiments. These experiments are designed in such a way to characterize the contributions of relay switching, equipment drift, and socketing. The socketing experiments, which dominates the overall repeatability error, includes periodical re-testing of a comparison IC throughout the testing of all ICs and instant

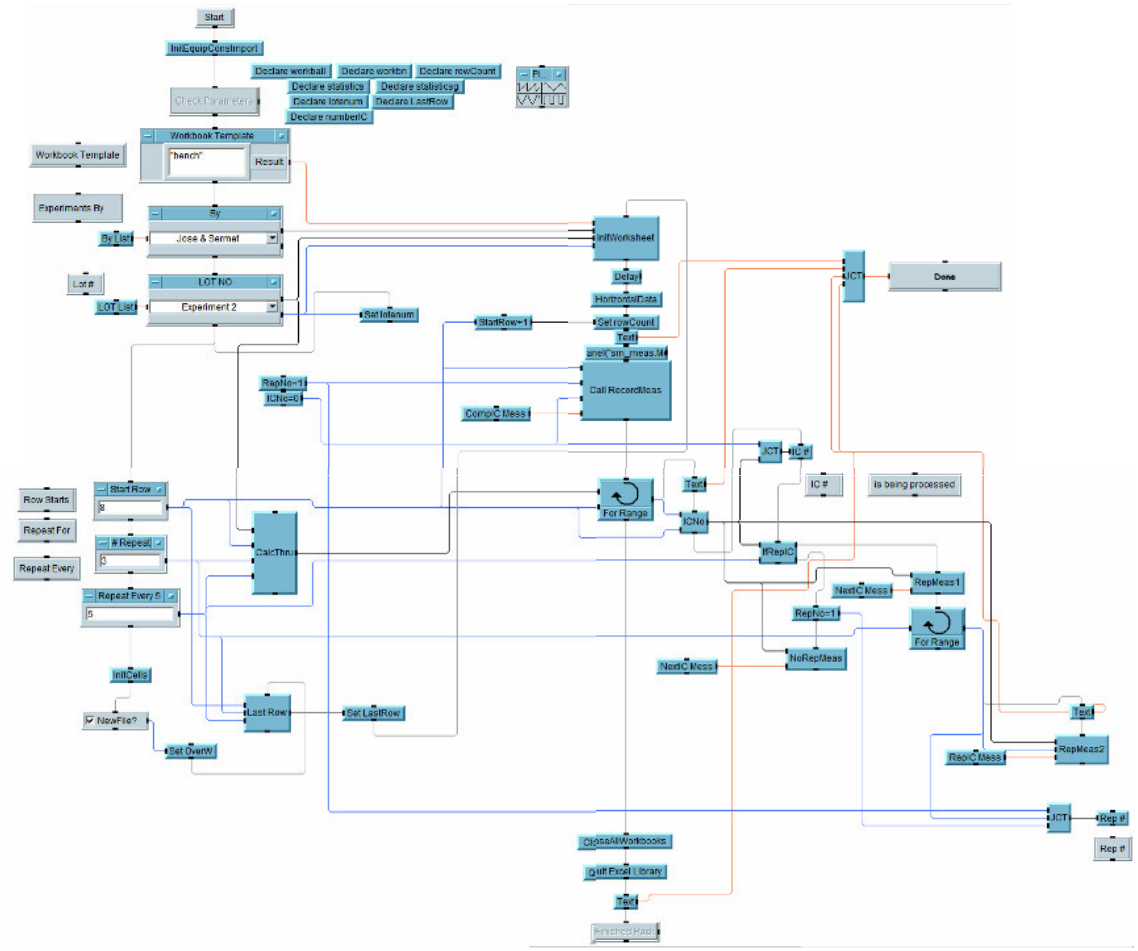


Figure 40: A snapshot of the Agilent VEE code.

re-socketing of specific ICs in each batch of 25.

3.5.2 Benchtop Alternate Test Experiment

The experiment starts with defining the constraints for the alternate stimulus optimization. These constraints also depend on the capabilities of the ATE, because the end-goal is to develop a benchtop alternate test which can be migrated to the ATE. Since the AWG is limited to 800Msps with 200MHz bandwidth, the at-speed test of the 850MHz DUT requires an up/downconversion scheme [43]. Figure 42 shows this setup: The center frequency of the optimized input stimulus is upconverted to 850MHz with a passive mixer driven by a free running LO. The upconversion mixer

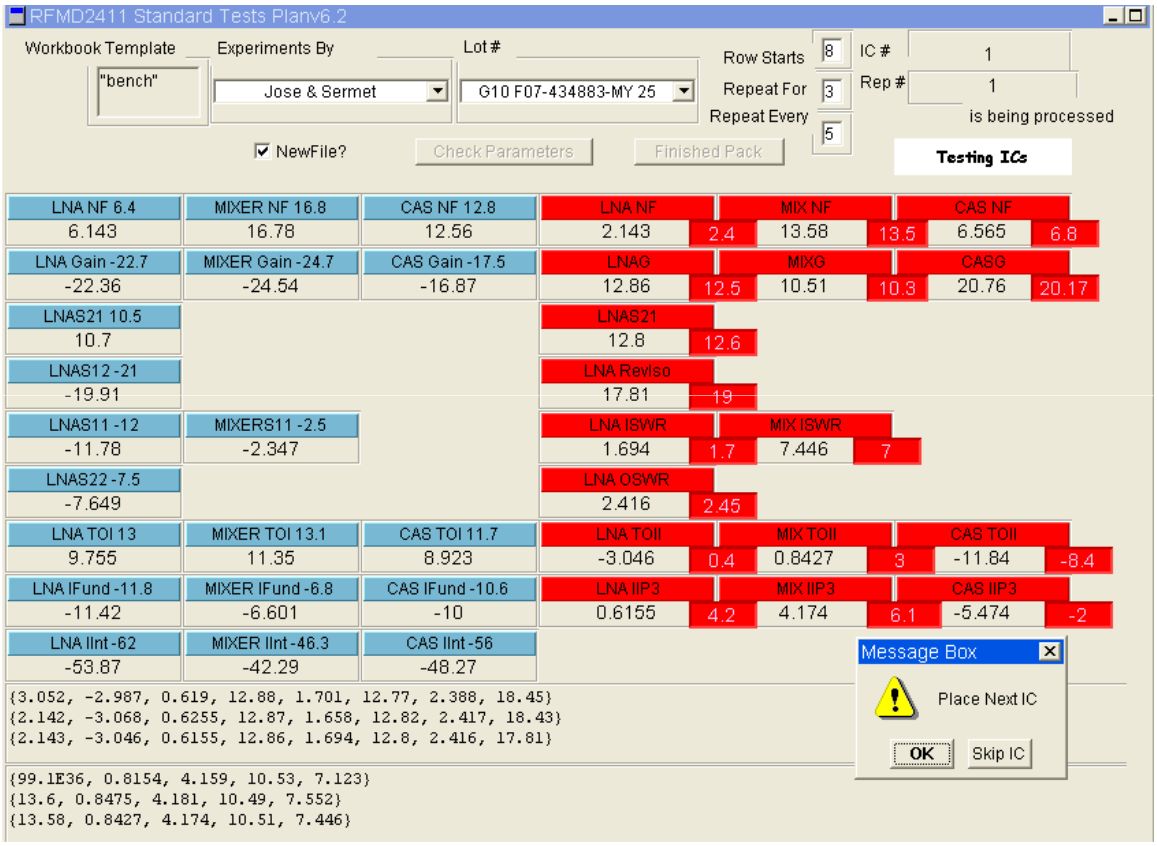


Figure 41: User interface for benchtop standard specification tests..

and the following image reject filter is characterized and embedded into the behavioral model. The response is downconverted to 50MHz with the mixer in the DUT driven by another free running LO. The response is passed through a low pass filter and logged by a sampling scope. After a 1024-point FFT, the magnitudes of twelve robust tones constitute the signature. This scheme is very similar to the simulation study performed in the previous section. *GAbS* driven by *AR* is used with *MARS* as learner and *FDART* engine for simulation. The search space for input tones is divided into eleven overlapping ranges with 20MHz bandwidth, Figure 43 shows these ranges, and gives an illustration of how GA progress through migration. Migration moves genes with high fitness values in one range to another range by transposing

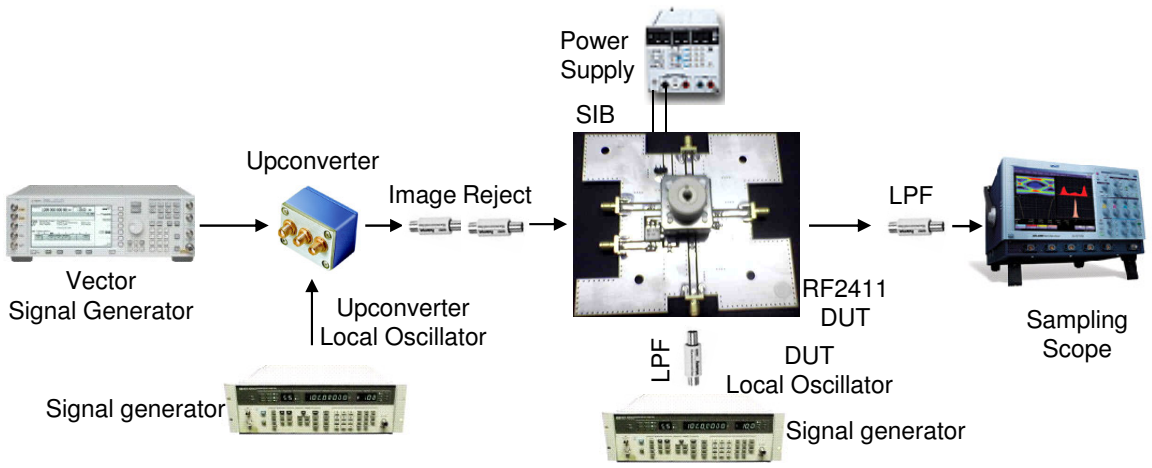


Figure 42: Benchtop alternate test setup.

the center frequency.

The top part of Figure 44 shows the optimized stimulus and a sample response signature for the benchtop alternate test. The stimulus is made up of seven tones around 177MHz ranging from −10 to −20dBm in amplitude as given in the top part of Table 3.5.2. This multitone signal is implemented on a vector signal generator using Agilent signal studio software [92]. All thirteen specifications, which represent LNA-only, mixer-only and cascade configurations, are predicted by this single stimulus applied to the cascade configuration.

Table 7: Alternate test stimulus tone locations and power levels.

Bench Frequency (MHz)	171	175	176	177	178	179	187
Bench Power (dBm)	−13.5	−15.5	−15	−10.5	−20	−19.5	−16.5
ATE Frequency (MHz)	135	136	137	138	139	140	141
ATE Power (dBm)	−12	−12	−12	−14.5	−12	−18.5	−19.5

3.5.3 Alternate Test on ATE

The ATE alternate test setup is identical to Figure 42, except that the vector signal generator and the sampling scope are replaced with the arbitrary waveform generator

and the digitizer on the ATE. Since it is a mixed-signal tester, the external LO sources are still required for up/downconversion. The integration of these LO sources is simple due to the fact that response signatures are composed only of FFT magnitudes, hence the LO sources are free running and require no synchronization [43].

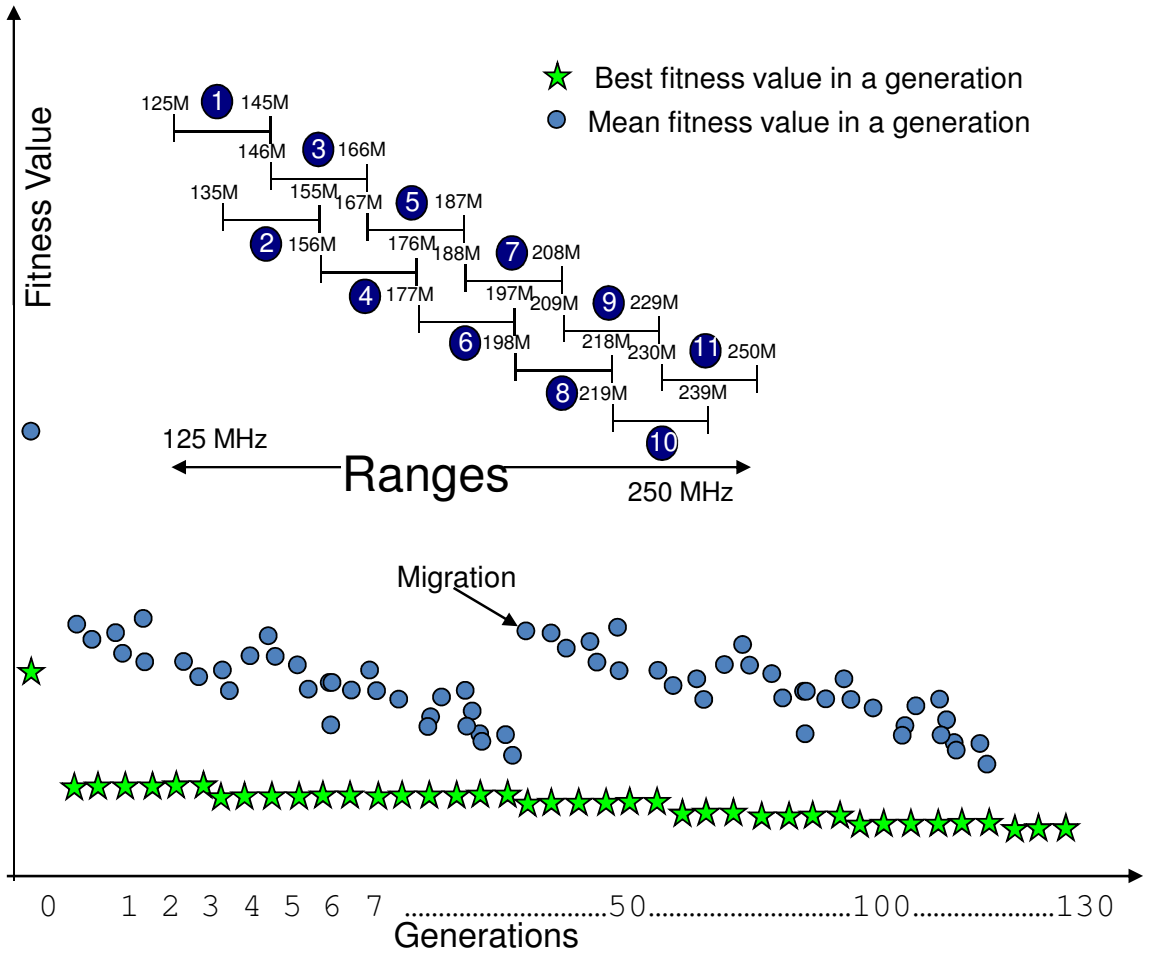


Figure 43: Illustration of GA search algorithm. The feasible frequency range is divide into eleven ranges. Migration moves genes with high fitness values in one range to another range.

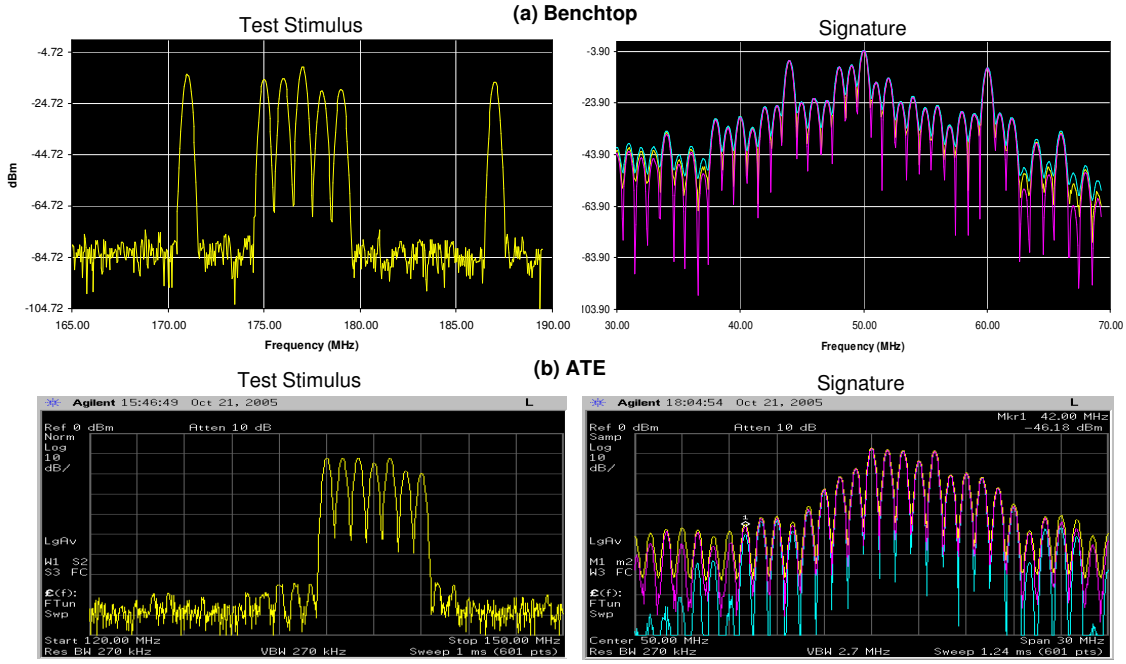


Figure 44: Alternate test stimulus and response for (a) benchtop setup, (b) ATE setup.

Originally, the alternate test stimulus optimized for the benchtop was to be used with the ATE setup. However, initial runs showed the AWG has an internal clock bleed filter, which generates a large roll-off for tones close to 200MHz. Also differential-to-single-ended conversion imposed an additional limit on the maximum amplitude of any tone. As a result, a filter representing clock bleed was added to the behavioral model, and the stimulus was re-optimized with the new constraints. Bottom part of Figure 44 shows the optimized stimulus and a sample response signature. The stimulus is made up of seven tones around 138MHz ranging from -12 to -20dBm in amplitude as listed in the bottom part of Table 3.5.2. As one can see, the additional constraints and the filter model have a significant effect on the optimized tones, which are further away from the 200MHz roll-off and smaller in amplitude when compared to the older stimulus.

3.5.4 Experiment Results and Comparison

The results from standard and alternate experiments show that alternate test time is an order of magnitude smaller than the standard method on benchtop. This is mainly due to the sequential nature of standard tests, for which the testing time almost linearly scales with the number of specifications. The setup times are dominated by the GPIB USB communications and add up close to 50 seconds for thirteen specifications; whereas in alternate test, there is only a single setup. Also, the test equipment for alternate test requires a signal source, a sampling scope and two LO sources, which overall costs \$144K, 60% less than standard test equipment used, which costs \$340K.

These benefits do not directly scale up for the ATE setup. Standard ATE setup times are far more optimized and the system cost of an ATE is already high regardless of its functionality. To make a healthy comparison, we used figures from a low-end commercial RF tester. This particular class of RF tester is comparable to the mixed-signal tester we have used in terms of settling time, yet still has the

minimal functionality to perform the required standard tests. We estimated the total standard specification testing time by multiplying the average per specification time with the number of specifications. For thirteen specifications, the proposed alternate test stimulus provided a significant 36% reduction in testing time. Furthermore, the mixed-signal tester together with two LO sources costs 48% less than the RF tester.

These cost benefits are significant only if alternate test is as accurate as the original specification test. In order to compare the accuracy, we use R^2 and AR. Figure 45 shows the ISO graphs for four sample specifications: LNA NF, mixer gain, cascade IIP3 and LNA OSWR. The x-axis shows the original specifications measured by standard tests while y-axis shows the specification values predicted by alternate test on the benchtop setup. The blue lines show the $\pm 3\sigma$ repeatability error for standard tests. The 7^{th} and 8^{th} rows of Table 3.5.4 list R^2 values for all specifications, for which LNA OSWR is the smallest with 74%. On the other hand, the ISO graph for LNA OSWR shows that all prediction errors are still smaller than the $\pm 3\sigma$ repeatability error shown in the 1^{st} row of the table. As a matter of fact, out of all devices in the validation set, only 3 devices in benchtop and 4 devices in ATE fallout of the $\pm 3\sigma$ limits for any specification. The 2^{nd} and 3^{rd} rows in the table list the maximum prediction errors for each specification, whereas 5^{th} and 6^{th} rows show the rms prediction errors. All rms prediction errors are smaller than the rms repeatability errors listed in the 4^{th} row. Hence, AR is very close to zero in terms of maximum errors and always negative in terms of rms errors. These results, together with high R^2 values, show that the implemented alternate test is as accurate as the original specification tests. Detailed graphs for benchtop and ATE prediction accuracies are shown in Figures 46 and 47 respectively.

In conclusion, alternate tests provide significant testing time savings and can be accomplished with low-cost testers. Although originally tailored for test operations with full access to netlist and process data, the proposed modification can be applied

to IP blocks with success. The ATE example shows 36% reduction in testing time, and 48% reduction in tester cost for a test case handled with behavioral models. Variations in implementation of the ATG methodology are also studied, and results show that there is no one-fits-all recipe.

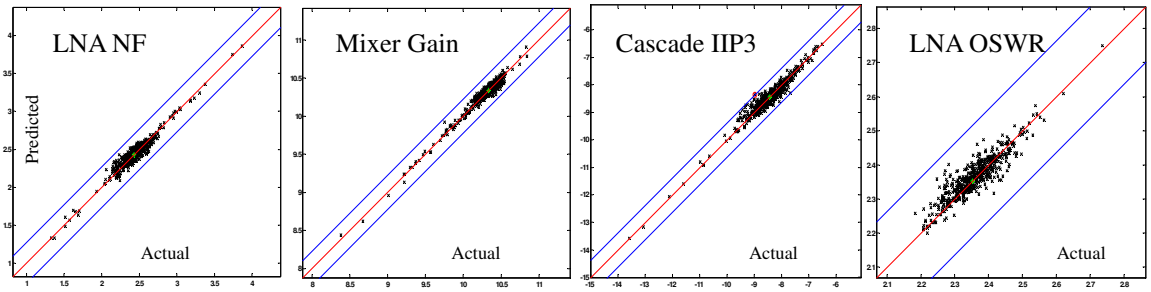


Figure 45: Actual versus predicted specification values for LNA noise figure, mixer gain, cascade IIP3 and LNA OSWR.

Table 8: Comparison of standard specification repeatability errors with alternate test prediction errors; (1) $\pm 3\sigma$ maximum repeatability error for standard, (2) maximum prediction error for alternate benchtop, (3) maximum prediction error for alternate ATE, (4) rms repeatability error for standard, (5) rms prediction error for alternate benchtop, (6) rms prediction error for alternate ATE, (7) R^2 correlation metric for alternate benchtop, (8) R^2 correlation metric for alternate ATE.

		LNA						Mixer				Cascade		
#	Error	NF	Gain	IIP3	ISWR	OSWR	RevIso	NF	Gain	IIP3	ISWR	NF	Gain	IIP3
1	Standard Spec $\pm 3\sigma$	0.27	0.29	0.62	0.11	0.16	1.94	0.33	0.23	0.82	0.60	0.48	0.42	0.62
2	Alternate Bench Max	0.20	0.27	0.54	0.13	0.12	1.27	0.28	0.17	0.82	0.45	0.40	0.33	0.64
3	Alternate ATE Max	0.35	0.26	0.54	0.07	0.12	1.37	0.23	0.24	0.66	0.35	0.54	0.69	0.99
4	Standard Spec Std	0.14	0.18	0.39	0.08	0.12	1.36	0.18	0.07	0.28	0.41	0.31	0.26	0.40
5	Alternate Bench Std	0.06	0.07	0.14	0.02	0.03	0.33	0.07	0.05	0.20	0.12	0.09	0.10	0.17
6	Alternate ATE Std	0.07	0.06	0.14	0.02	0.04	0.33	0.06	0.06	0.17	0.10	0.09	0.14	0.22
7	Alternate Bench R^2	94%	88%	93%	88%	85%	85%	79%	97%	97%	85%	86%	95%	94%
8	Alternate ATE R^2	94%	87%	94%	88%	74%	83%	85%	93%	95%	92%	86%	89%	92%

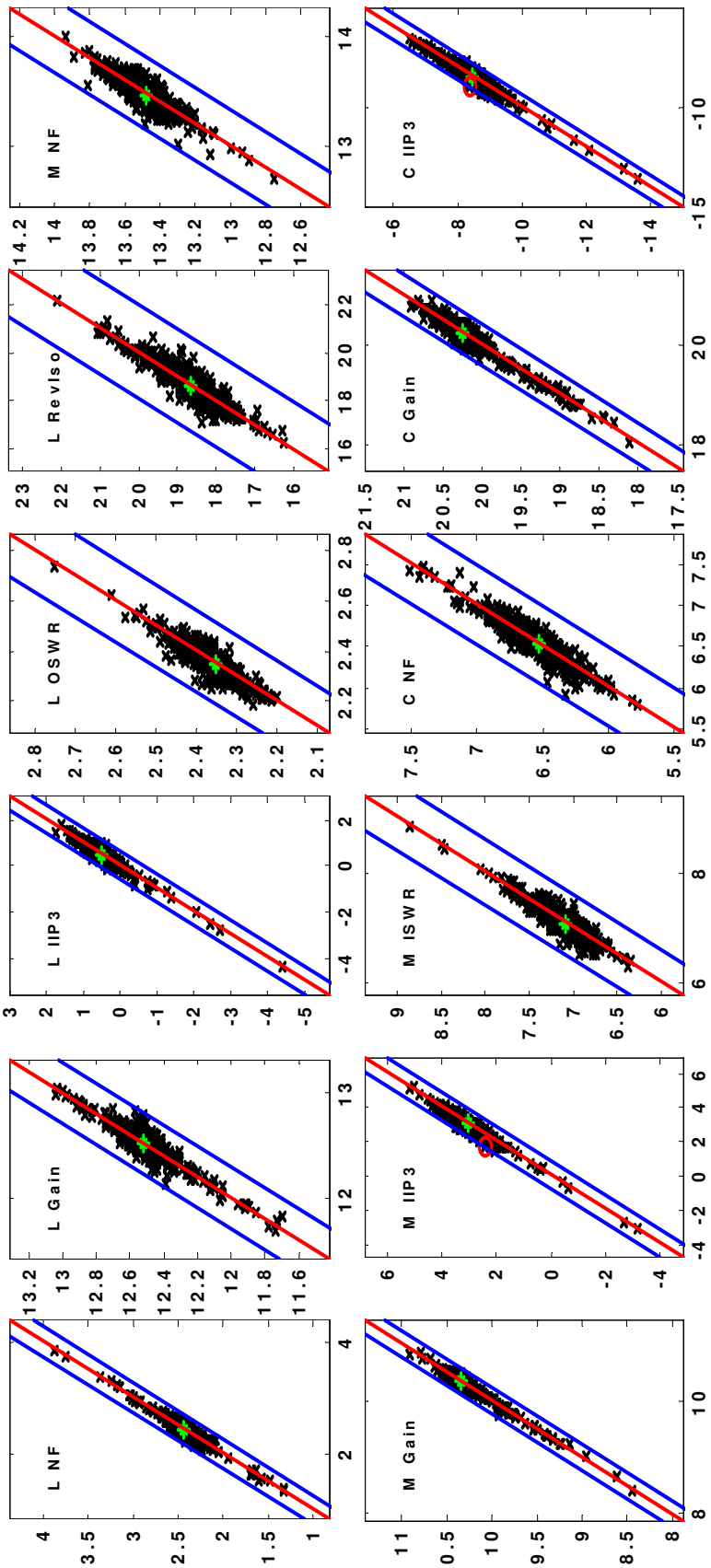


Figure 46: Actual versus predicted specification values for all specifications on the benchtop setup.

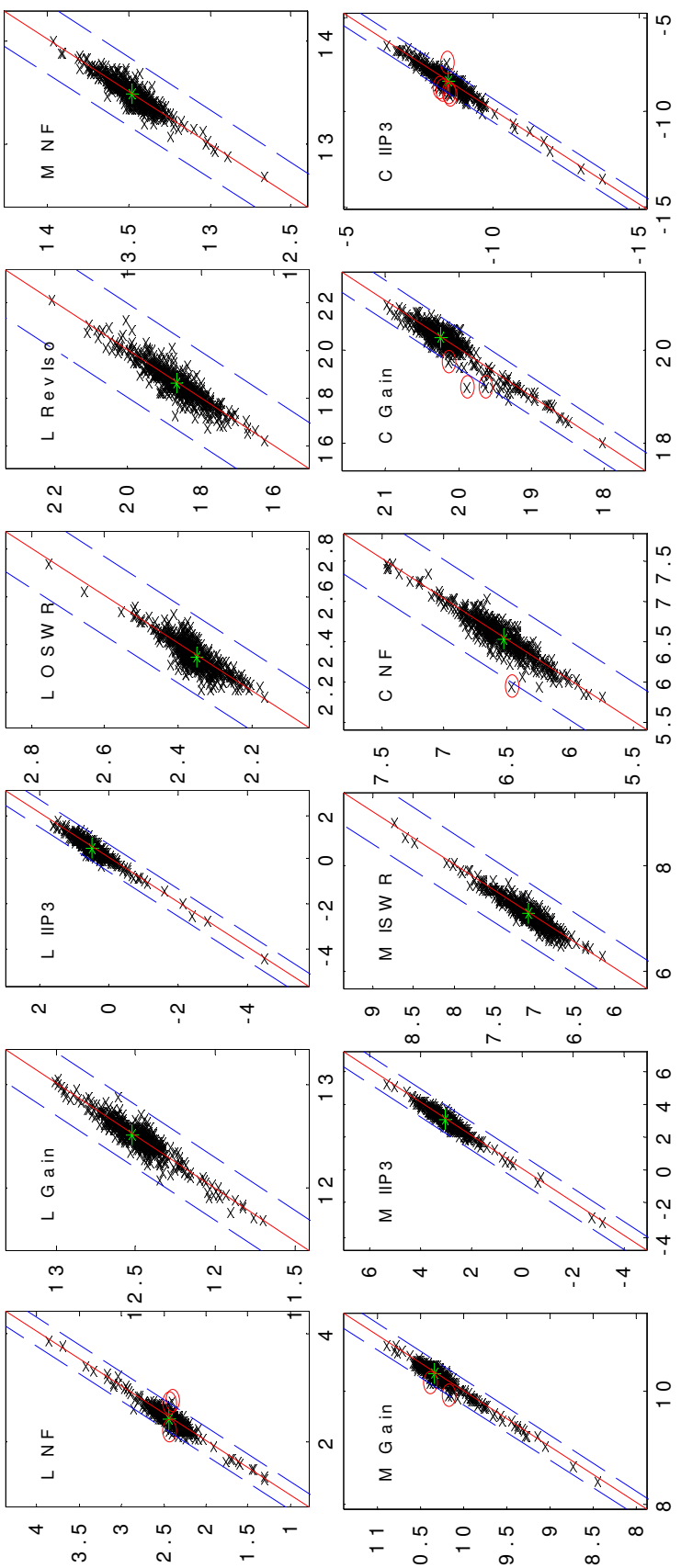


Figure 47: Actual versus predicted specification values for all specifications on the ATE setup.

Chapter IV

FEATURE EXTRACTORS FOR BUILT-IN TEST

In this chapter, methodologies for the BIT of RF devices is presented. Extension of alternate test methodology with DfT features, such as built-in and built-off alternate testing, becomes critical for highly integrated systems. SOPs, being on the extreme end of this spectrum, favor built-in alternate testing, taking advantage of the flexibility provided under limited observability and controllability. In this scheme, the BIT makes use of additional hardware components integrated into the system under test at the package or chip level. These hardware components correlate the high-bandwidth test response to low-bandwidth signatures, while extracting the test-critical features of the DUT. The scheme makes use of supervised learning [93, 94], as was described in the other alternate test techniques, to map these extracted features into the SUTs, which otherwise are too complicated to decipher by plain mathematical analysis.

At low frequencies, alternate testing is based on sampling the test response using an ADC and analyzing the digitized response in the external tester. In order to use alternate testing at frequencies in the multi-GHz range, the test waveforms need to be very simple and the evaluation of the test response should be handled by practical hardware-based test response feature extractors.

The call for a testable SOP results in a conflict of interest between the degree of integration afforded by the design process and the level of testability achievable by an external tester. A viable solution is to place the ATE functionalities in close proximity to the SOP module to be tested. This improves the test-access speed, minimizes test signal degradation, and increases controllability and observability of the signals internal to the DUT. One such candidate is the loadboard itself, where the

test functions are migrated from the external tester to the additional circuitry built around the system under test. The additional circuitry retains the ability to apply a high-speed stimulus to the system under test and capture the high-speed test response, which otherwise are degraded by the cable parasitics of the low-bandwidth external ATE. The solution, BOT, presents a low-cost alternative to the prohibitive cost of a classical ATE. The other alternative, BIT, pushes the external tester functionality into the package and even into the bare dice wherever possible and is consequently a much more aggressive version of BOT.

Note that without BOT or BIT, high performance SOPs may not be economically testable. This is because the cost of external test equipment for test signal speeds in excess of 1GHz is a very significant cost adder. However, multi-gigahertz system designs are now becoming quite routine for high-bandwidth communications. The test economics is greatly improved by having high-speed test functions on the loadboard (BOT) or the SOP itself (BIT) augmented with low-bandwidth communication with a low-speed external tester. This allows high-speed systems to be tested with a low-cost external tester without loss of test quality.

4.1 RF Test Strategies with DfT Features

While migrating external tester functions to the proximity of the DUT, there are two different possibilities: (i) the DUT is considered as an end product without having dedicated test functionality internal to the device and, hence, the test support circuitry is built around the device, or (ii) the test support functions are implemented within the device as an integral part of it. The first approach, BOT, is suitable for applications where the internal design of the DUT cannot be modified for test purposes and the package itself does not constrain the speed of the test signals that can be applied to the DUT. The second approach, BIT, is more of a DfT methodology. The support functions are implemented within the same package or even in

the same chip area. In this approach, the device is modified to incorporate some additional functions within the chip by using dedicated test circuitry [95–98] and by reusing components [57] such as ADCs and digital-to-analog converters (DACs) already available at the system level. The introduction of test circuitry into the device may violate original design constraints, for example, device matching and parasitic loading, and as a result additional design iterations may be needed during system design. Consequently, BIT is feasible only when it can be integrated into the system design flow.

Irrespective of whether BIT or BOT is used, the loadboard is a necessary component in a production test environment and typically routes the signal from/to the test-head of the external ATE to/from the DUT. Figure 48 shows the role of a loadboard in a high-end conventional ATE environment. In this environment, the loadboard contains a low-parasitic socket to hold the DUT, power and ground planes, signal traces, and switches and relays to multiplex external tester resources. The external tester generates the entire test stimulus, and the DUT response is directly relayed to it. High-bandwidth data transfer is performed at the operation speed of the DUT. In the BOT, this channel is replaced by a low-bandwidth connection, which is utilized to send test control signals as well as a low-speed test stimulus and also to receive compact signatures extracted from the DUT response. The high-speed test stimulus and response signature generation is handled at the loadboard by means of customized signal generators, samplers, converters, modulators, demodulators, multiplexers, and demultiplexers. Modems convert the low-speed stimulus coming from the external tester into high-speed stimulus required by the DUT; similarly the response is down-converted.

Alternatively, the test stimulus can be generated on the board and the response can be compressed into a signature by samplers and converters. Figure 49 depicts a general BOT strategy. BOT implements complex test signal generation and test

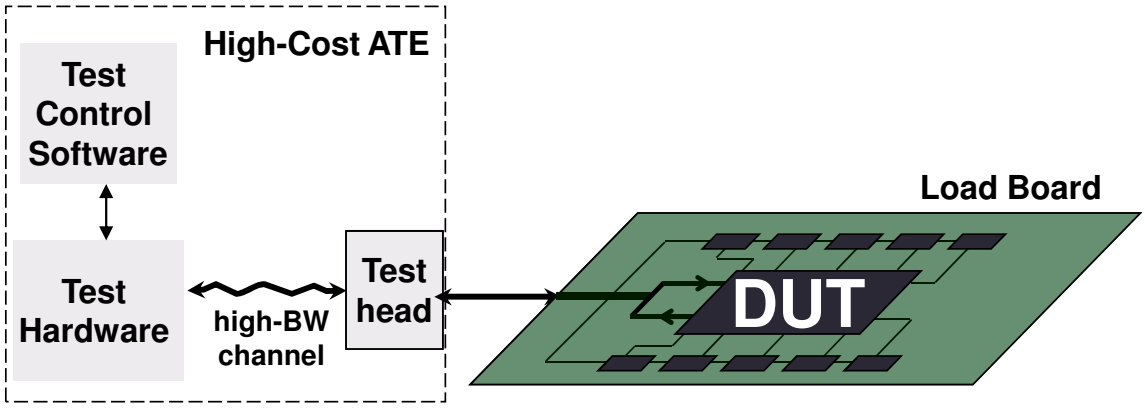


Figure 48: Loadboard in a high-end conventional ATE environment.

signal modulation schemes without employing expensive "feature-enriched" testers at the expense of higher loadboard manufacturing cost in production testing. The high-speed test signal processing is all done on the loadboard itself under external tester control. The tester employed is typically low-cost with low-speed digital and analog test data transfer to and from the loadboard.

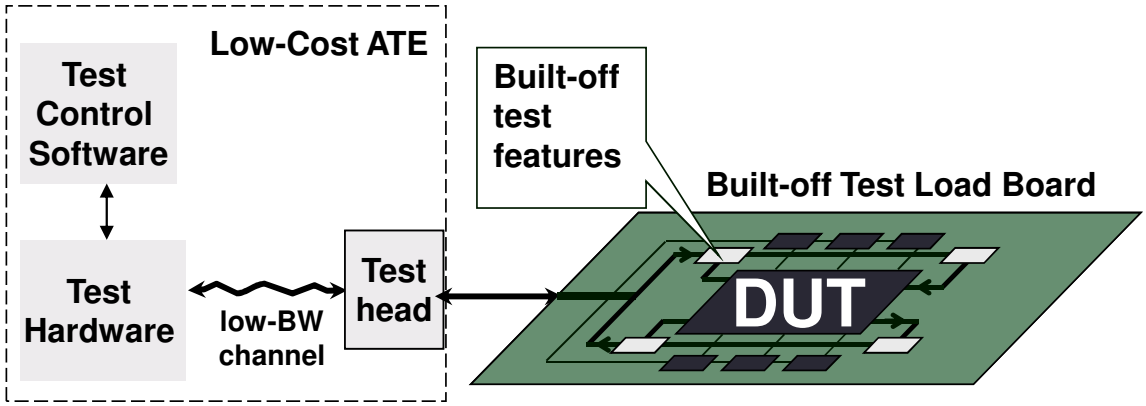


Figure 49: General BOT strategy.

BIT, as shown in Figure 50, pushes the tester functions into the DUT in order to overcome the two main challenges in testing, excitation of the DUT and propagation of the response to an external test node. As the complexity and integration of SOPs increase, both issues become harder to tackle, and the test paradigm shifts to solutions where DfT [99–101] is employed to improve the controllability and observability of

internal nodes [102, 103]. The IEEE 1149.1 (JTAG) [35] boundary scan standard provides an effective means for test access to internal modules of the DUT [104] for testing static faults in digital ICs, which are faults that can be sensitized with a single operation, such as stuck-at or flip; however, its JTAG counterpart in mixed-signal testing, the IEEE 1149.4 standard [37] is limited by its low-bandwidth [105]. Hence, built-in testing of analog, RF, and mixed-signal electronics still presents the following major challenges:

- On-chip generation of high-speed test stimulus using low-cost hardware.
- High-speed on-chip response acquisition followed by analysis or response compaction.

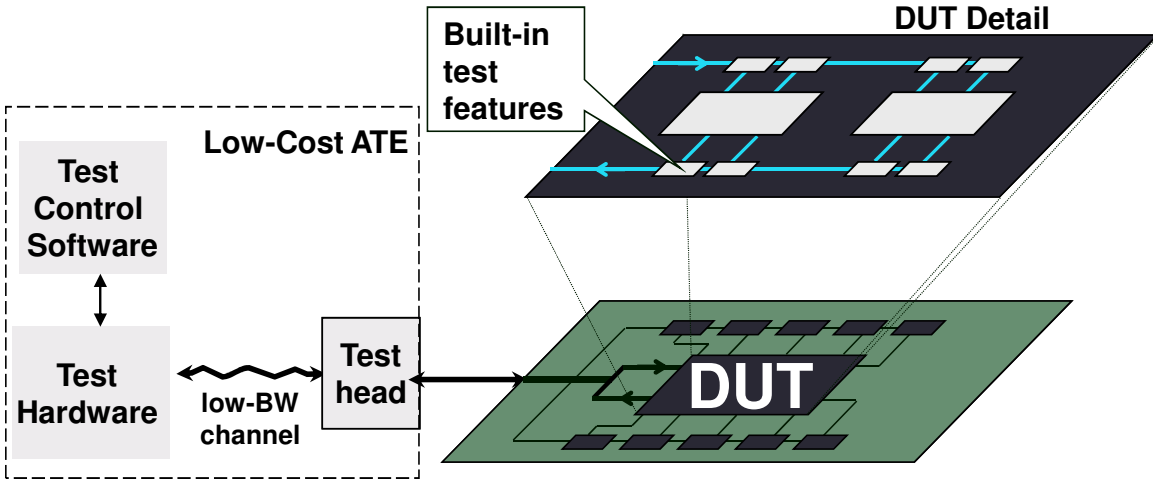


Figure 50: General BIT strategy

In BIT, low-speed communication takes place between the external tester and the built-in circuitry inside the DUT. This media is used to start/stop a test or run status checking commands, while the BIT circuitry performs the rest of the testing in situ. Although this approach addresses the tester cost and test access limitation problems, the large chip-area taken by these circuits, especially in mixed-signal testing, makes

it often uneconomical for testing all chip functionalities in situ. With the evolution of highly integrated systems such as SOPs, this area overhead is less of a concern thanks to the reuse of already embedded components such as DACs, ADCs, and on-chip DSP.

The embedded functions in BOT and BIT carry different levels of intelligence. They can be implemented in such a way that they create all necessary test vectors and analyze the DUT response on demand, generating a conclusive result about the state of the device. The resulting approaches, built-off self-test (BOST) and BIST, are complete and independent of any external tester help; however, they may require high processing power especially when analog and RF components are to be tested. Such components are more likely to benefit from a low-speed, low-pin-count external tester, which analyzes the response signature and generates the test control and low-speed excitation signals. In this kind of "less intelligent" support, the external tester can also be utilized to test the operation of BOT or BIT components before testing the DUT; this scheme provides flexibility when additional tests are required in the production line. On the other hand, a true "self-test" is not limited to the production line, since it can be applied throughout the lifetime of the device periodically or right before it is turned on. This may be an important criterion for critical systems that are likely to deteriorate over time, such as in space applications. Such schemes are more likely to be implemented as BIT since BOT requires a loadboard.

In a traditional production test approach for testing of analog and mixed-signal circuits, the functional specifications are measured using the appropriate tester resources and using the same kind of test stimuli and configuration, with respect to which the specification is defined [106]. For example, a multitone signal generator is used for measuring distortion, or a ramp generator is used for measuring integral nonlinearity (INL) and differential nonlinearity (DNL) of ADCs and DACs. The measurement procedures are in agreement with the general intuition of how the module behaves, and, hence, the results of the measurement are easy to interpret, in contrast

to the concept of "alternate test". Since this is as direct as it gets to measuring a performance metric, it is called "direct measurement of a specification".

In a direct measurement approach using BIT, the external ATE functionality is designed inside the DUT for applying appropriate test stimuli and measuring the test response corresponding to the specification. In [107], adjustable delay generators and counters are implemented next to the feedback path of a phase-locked loop (PLL) to measure the rms jitter. Since the additional circuitry does not modify the operation of the PLL, the same BIT circuitry can be employed on-line. Reference [107] also discusses different ways to measure properties like loop gain, capture range, and lock-in time by modifying the feedback path to implement dedicated phase delay circuitry. All these built-in test components are automatically synthesized using the digital libraries available in the manufacturing process. This kind of automation provides scalability and easy migration to different technologies. The approach of [108] is similar in the sense that the extra tester circuitry is all digital and can be easily integrated into an IEEE 1149.1 interface. The built-in test reuses the charge pump and the divide-by-N counter of the PLL in order to generate a defect oriented test approach, which can structurally verify the PLL. While [107] can also be implemented as a BOT, [108] is limited to BIT, because a multiplexer must be inserted into the delay sensitive path between the phase detector and the charge pump. Since both examples employ all digital test circuitry, their application is limited to only a few analog components like PLLs, where digital control is possible.

The works of [57, 109–111] attempt to implement simple on-chip signal-generators and on-chip test response data capture techniques for testing the performance of high-frequency analog circuits. The communication between the BIT hardware and external world takes place through a low-frequency digital channel. In particular, [110] measures the spectral content of the test response using direct down conversion of RF test stimuli and test response waveforms. Although the chip-area taken by additional

test circuitry is still a concern, it shows the feasibility of using BIT for measuring performance of high-frequency embedded analog/RF blocks in-situ.

With regard to BOT, direct measurement techniques for different classes of analog circuits are discussed in [3]. The circuitry for measuring one test specification is reconfigured to measure another using a set of relays and switches on the DUT loadboard. Typically, the loadboard test circuitry is designed with the DUT designer's input, unlike the method presented in [112], and takes several weeks to debug.

Although the direct measurement procedures are conceptually simple, this approach has inherent drawbacks as described below:

- Multiple specification measurements require different kind of resources, which are difficult to build either on-chip or on the loadboard due to high area overhead.
- A longer overall testing time is required since measurement of multiple specifications cannot be performed simultaneously.

As a result, direct measurement techniques are not sustainable solutions for BIT as test resource requirements are high, associated BIT hardware overhead costs are prohibitive, and the time necessary for testing each specification separately increases the overall manufacturing cost.

4.2 RF Alternate Testing with DfT Features

4.2.1 Alternate BIT and BOT Examples

Recent literature addresses applications of alternate testing to RF components with DfT features. In [43], the loadboard modulates the baseband test stimulus provided by a low-cost tester and uses the resultant RF signal to stimulate a LNA. The response is down-converted on the loadboard and lowpass-filtered to generate a signature that can be transferred to the tester through a low-bandwidth channel and analyzed using

alternate test principles. The application follows the generic modulator based BOT scheme in Figure 51. An alternative to this scheme is using a simple signal generator that can be implemented on the loadboard. In Section 3.1 of this thesis, we described an alternate test generation methodology that seeks an optimal superposition of sinusoids. Simulation results suggest a single sinusoidal, which has two orders of magnitude smaller frequency than the nominal and can be used to excite an RF LNA. The response can be sampled with loadboard capabilities. A different alternate BOT approach was discussed in Section 3.2, which employs the bias control voltage of an RF power amplifier as the test stimulus. The use of current measurements in response acquisition is proposed as a non-invasive alternative to voltage measurements in RF applications where tapping into sensitive nodes is prohibitive.

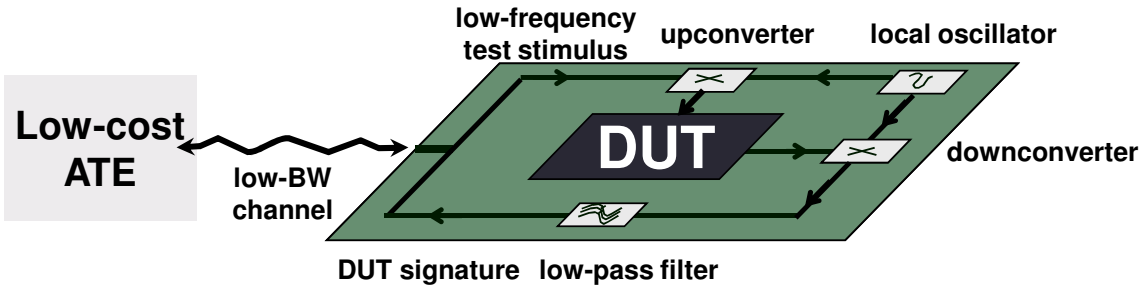


Figure 51: Modulator-demodulator-based alternate BOT scheme.

In [65], another version of the BIT scheme is proposed, which deviates from the self-test paradigm in order to minimize additional test-related hardware placed inside the chip and to reuse the existing test hardware already present for testing the digital section of the system IC. Unlike BIT schemes discussed in [96–98], the DUT response is analyzed externally inside a low-cost ATE. Since the DUT test response waveform is transformed to a digital bit stream and scanned out through the scan chains of the digital cores, the approach can be integrated with an IEEE 1149.1 based scan structure. Hence, the proposed technique attempts to solve the limited test access

problem for embedded analog modules in system ICs to a large extent and can be used for testing the embedded passives at the assembled level as opposed to the substrate level. The test response waveform is reconstructed for analysis in the external tester, and from the reconstructed test response waveform, the DUT's specifications are predicted using the regression analysis discussed in alternate test.

In another kind of BIT approach, the circuit topology is changed using additional circuit elements to make the circuit behave differently from what it is designed for, and this modified functionality is usually easy to measure in the production test environment. The catastrophic faults that make the original circuit performance fail also causes the reconfigured circuit performance to deviate. The latter performance deviation is measured during the production test, and pass/fail decisions for the original circuit are made. Oscillation-based tests (OBT) [113–115] act on the above principle, which reconfigures analog filter circuits into oscillators using additional feedback components. This BIT technique detects catastrophic faults in DUT by measuring the deviation in oscillation frequency and amplitude. In recent years, the above defect-oriented BIT technique has been integrated with the regression modeling approach commonly used in alternate tests, and the modified OBT is used for predicting the specification of DUT under parametric failure conditions. The modified technique relies on the fact that the original circuit and the reconfigured circuit share almost all the circuit components, hence, a direct correlation between the original circuit performance and the modified circuit performance (the latter performance is not a design goal) values can be established when the circuit parameters vary. This correlation is computed using circuit simulation under parametric variations using regression analysis previously used in alternate testing. The modified OBT, referred to as predictive oscillation based testing (POBT) [116], predicts the performance of the original circuit by computing the above correlation and measuring the oscillation frequency of the modified circuit during test. One inherent drawback in OBT approaches is that

very few circuits other than analog filters can be reconfigured into oscillators.

A built-in response acquisition presents a significant problem in testing RF submodules. In mixed-signal environments with built-in ADCs, the analog response can be fed into the ADC and scanned out to the external tester in digital form after compaction. However, in RF systems, the inherent ADCs are configured to process nearbaseband signals, so their performance is not adequate to process high-frequency passband responses. In [117, 118], this problem is tackled by introducing a statistical sampler that compares the analog response with noise. The power spectral density (PSD) of the resultant digital bit stream is a representation of the original PSD with an increased noise floor. In this chapter, we extend this methodology by an automatic feature extraction scheme that detects the PSD components above the noise floor introduced by the statistical comparator, and uses these components with a nonlinear mapping model to predict device specifications like gain, IIP3, NF and PSRR. This scheme presents an extension to the alternate test methodology, in the sense that the scheme can compensate imperfect tester conditions simulated with a random fluctuation superposed on to the ideal input stimulus.

4.2.2 Direct Measurement versus Alternate Testing

SOP requires non-orthodox test methodologies that can keep up with the test access problems amplified by the inherent integration, and do away with the prohibitive cost of high-end external testers. BOT and BIT strategies propose a solution to these problems by placing high-bandwidth test access either next to or within the package. The SOP test challenge also calls for automated test solutions, which are not only generic enough to cut down custom test support development cost, but also customizable enough to make the test-support financially feasible. This requirement ensures that the turnaround time associated with test generation and test hardware development is minimally reflected in the device manufacturing cost. Although different

direct measurement based test approaches reviewed in Section 4.1 propose promising results for stand-alone devices, their application at a system scale is not feasible because of the need for custom test-support hardware for every embedded module to be tested. Since they do not provide a generic methodology to handle direct measurement of different specifications, each specification to be tested increases the overall turnaround time for product development, as well as increasing the test area overhead and testing time for every device.

Alternate test methodologies propose generic solutions for embedded analog and RF components, which cover a large range of system components available in SOP, namely, embedded passives, opamps, filters, LNAs, mixers, power amplifiers, and others. The ability to predict multiple specifications using a single test reduces the test hardware complicity, area overhead, and testing time. Furthermore, statistical sampler-based extensions are compatible with applications utilizing digital scan architecture (IEEE 1149.1), since the resultant bit stream can be relayed to the digital signal processors in the package at no additional cost.

Since alternate BOT methodologies propose a systematic way to handle a large range of specifications and submodules, their integration into the product flow does not increase the complexity and cost significantly. Although the loadboard will be populated with extra components to accommodate test-related signal processors, the increase in board design time can be compensated with automation already present in the traditional loadboard manufacturing flow [112]. The manufacturing cost of traditional loadboards is dominated by the quality of the material, the many levels of power planes provided, and the necessity to use only "golden" boards, boards that very closely follow the specifications of the original board design. In the case of BOT, the extra cost of signal processor ICs, their routing and assembly will not be significant when compared to the traditional loadboard figures. Furthermore, the use of BOT will benefit from low-end ATEs, which provide two orders of cost reduction compared

to high-end ATEs necessary for traditional tests [119–121]. This reduction is still one order greater than the manufacturing cost of many typical complex loadboards. The only practical limit for loadboard complexity is the fixed area dictated by the interface of the production testing equipment. It is important that when it comes to testing of very complex systems like SOPs, one of the main problems is feasibility rather than cost [122, 123].

A viable SOP strategy using alternate tests should generate specification-oriented tests considering only the component specifications that progressively develop a violation at the system level. The first step will be analyzing the system specifications to break them down into related component specifications. This process is usually a part of the system design; hence, it will not induce further effort. Then, all related specifications can be tested by a single alternate test per component. Some system-level specifications that cannot be verified by a collection of individual component performances will further be covered by system-level alternate tests. Reference [28] presents an example to this scheme, where system-level specifications of the RF subsystem of a narrow-band wireless transceiver are verified by alternate testing generated on high-level models of the system. In Section 3.3 of this thesis, we presented a comprehensive test generation scheme with behavioral models. High-level modeling speeds up the simulation intensive features of alternate testing, which are not feasible for SOPs at the netlist level. Furthermore, the inherent complexity of SOPs makes built-in approaches more favorable than BOT solutions. On the other hand, a joint built-in/built-off approach can add more value to the package area, where module-level access and DSP is handled by built-in tester components, and more area-intensive tester functions like analog signal generators and modulators-demodulators are migrated onto the loadboard.

4.2.3 Alternate Test Development Flow with DfT Features

Alternate testing provides a framework for high-volume manufacturing tests of components and systems that are evaluated by analog specification boundaries. This framework takes many diverse forms in implementation depending on test benchmarks and specific requirements of the DUT. Although it is not possible to cover all different implementations in this section, different applications share a common flow with greater emphasis on different stages. Figure 52 shows this generalized alternate test development flow and its HVM application.

First, the optimization space is defined by the stimulus range and the available measurement equipment. Possible DFT features are also considered at this stage if it is still possible to make an impact on the design. Second, optimization models are created. These models may be netlist-level if the IP is available and the simulation time is not a bottleneck. However, for complex systems such as SOP, only high-level models are feasible. These models need to capture enough information to represent expected process variations and facilitate efficient simulation for test optimization. Once these models are created, a sample set with process variations is generated by MC analysis or statistical design of experiments (DOE). In the fourth step, the stimulus, signature extraction, and, if applicable, DfT features are co-optimized to fit the test envelope defined by test cost, test time, and test quality metrics.

The optimization step is usually iterative and makes use of techniques described in the previous chapter, such as genetic algorithms or response surface methods that are more suitable for complicated, nonlinear, partially continuous domains typical for test generation. If the optimization does not converge, or it converges to a suboptimal metric as measured in the fifth step, then the models may be tweaked to better capture the problematic response variables, and the sample set can be extended. Steps two through five are iterated until an alternate solution composed of stimulus, measurement equipment, response signature generation algorithms, and possible DFT

features are delivered to satisfy the goal metric. In step seven, both classical specification tests and the optimized alternate tests are performed on a sample set of DUT hardware. Next, correlation models are generated. These models map response signatures from alternate testing to the actual specification values measured by classical tests. In HVM, right side of the figure, alternate stimulus is applied to DUT, and the measurements are converted into a response signature. Then the mapping functions from step eight are used to predict specification values. These values are tested against pass/fail limits unless the response does not fall into the expected envelope defined by the training set in step six. The result is either a pass/fail decision, or the DUT is marked as an outlier. Outliers may be kicked back to classical specification testing at the expense of testing time and tester cost. Once the number of outliers exceeds a certain limit, the mapping functions are re-calibrated with an improved sample set including these elements.

4.3 Feature Extraction with a Noise Reference

The first approach demonstrates a BIT scheme for RF components embedded in a system with available DSP resources. This scheme addresses the precise analog signal generation and response acquisition problems associated with previous analog BIT solutions. Automatic extraction of test response features from the DUT response makes this scheme favorable for DfT flow. The DUT is excited by an embedded oscillator, and a one-bit noise-referenced comparator captures the response. The resultant bit stream is fed into the digital scan chain of the system and recollected at the DSP resources, which reconstructs an approximation of the original spectral response. The proposed algorithms extract features from this spectrum, and they are mapped into predictions for RF SUTs using alternate test principles. Two key features of the proposed approach are: (i) compensation for imperfect stimulus and (ii) use of a low-speed low-resolution noise-referenced comparator for sampling. These

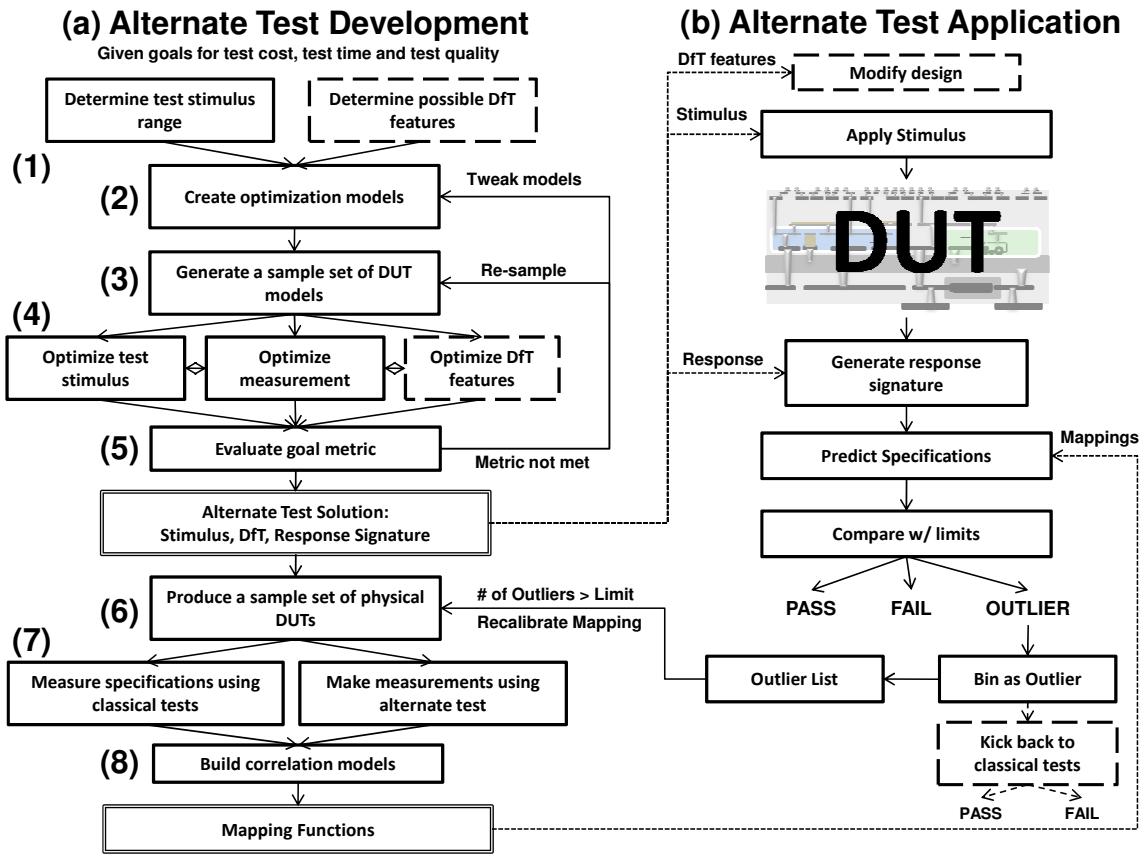


Figure 52: (a) Alternate test development flow, (b) alternate test application.

two features make the proposed scheme ideal for low-cost and low-area overhead BIT implementations.

Note that the second aspect above was first proposed by Negreiros, Carro and Susin [124]. However, while [124] presents a generic approach to noise-referenced fault detection, no bridge to traditional specification-based analog/RF testing was developed. Fault-based testing has its own caveats, namely the lack of accepted fault models and failure mechanisms. For the technique of [124] to succeed, it is desirable to port it to the area of specification-based testing, which is essentially fault-independent. In other words, it should be possible to infer the analog/RF specifications of the DUT directly from the test response. Such a step is not possible without proper response feature extraction as discussed in this section. It is shown

that once feature extraction is performed, it is possible to determine all the DUT specifications from the observed digital test response very accurately. In addition, the generation of the demonstrated BIST is fully automated; hence it can be integrated to the design flow as a designer-friendly DfT step. The methodology can also be extended to implement BIT schemes for systems that lack the power of DSP resources, but instead have access to a low-cost low-speed digital tester.

4.3.1 Noise-Referenced Alternate Testing Methodology

In this section, specification-based alternate testing methodology is expanded in a way to accomplish low-cost built-in self-test of RF components. The methodology follows the alternate testing theory described in previous sections, but uses a novel way to extract features from response signatures. Previous work uses samples of the signature directly, but that approach experiences limitations in terms of sampling speed and timing accuracy when applied to very high-speed RF components. Use of indirect features relaxes the constraints on sampling and accuracy, hence implements a method suitable for BIST applications.

The sampling of the signature is a critical part of alternate testing such that the speed and accuracy of sampling mostly defines the accuracy of predictions. For RF components operating in the gigahertz range, this requirement defines a problem, since the Nyquist sample rate of such signals and their harmonics may far exceed the capabilities of ADCs already present on the system. Even if such ADCs are present on the system, they introduce significant area overhead and signal degradation when interconnected as a part of the BIST scheme. Instead, [124] introduces a low-cost sampler with low-area overhead. In this scheme, the DUT output is compared with noise to generate one-bit output such that the PSD of the resultant bit stream can be processed to differentiate between fault free and faulty circuits. The comparison process can be modeled as a hard-limiter and when the input $x(t)$ is a stationary

process with zero-mean, the autocorrelation of the output $y(t)$ is given by [125]:

$$R_y(\tau) = \frac{2}{\pi} \arcsin(\frac{R_x(\tau)}{R_x(0)}.) \tag{10}$$

Equation 10 states that the statistics of the input will be preserved at the output of a hard limiter. When $x(t)$ is compared with white noise, the resultant autocorrelation will be a scaled and biased form of the original. The level of bias depends on the amplitude of white noise [126], which must be greater than or equal to the amplitude of signal to be compared. Finally, the result will also be transformed by the arcsine function. Under reasonable conditions, the nonlinear regression mapping can trespass the effects of all three operations. From a practical point of view, the noise comparison process is analogous to a random sampling of the signal; hence, as the above discussion suggests, the spectral content at the output of the comparator is identical to a scaled version of the original plus some noise floor resulting from random sampling. Main spectral components of this signal may still be above the noise floor, and carry valuable information about the original signature. In this section, we implement a feature extractor, which performs a fast fourier transform (FFT) operation and then automatically detects spectral components above the noise floor introduced by the noise reference.

The feature extractor is based on wavelet transformations. Coiflets of the second order [127] are used to decompose the original spectrum (s) into eight levels of coefficients, and then some of the coefficients are eliminated using a soft minimax de-noising method [128]. The signal is reconstructed by the remaining coefficients resulting in the de-noised signal (D_s). The residual (R_s):

$$R_s(f) = D_s(f) - s(f) \tag{11}$$

defines the frequency-dependent noise floor. After applying a guard band above this level, all the spectral components below this floor are removed. The remaining

spectral components define the final features extracted for mapping. The remaining steps follow classical alternate testing methodology; a set of training examples are simulated with process variations given by process parameter distributions, and MC simulations define the feature extractor parameters and the mapping between these features and specification values.

Figure 53 depicts the overall testing process: the response signature is compared with the noise reference and one bit of data is generated at a time. A large number of these bits are recollected at the feature extractor and the spectral results are trimmed accordingly after the FFT. The few resultant spectral elements are fed into the mapping model to generate predictions for SUTs. These specifications are compared with hard-coded threshold values and the test process displays a final bit representing either pass or fail. The DSP components of this process are summarized in Figure 54 together with the algorithm to generate feature extractor parameters.

4.3.2 Noise-Referenced Alternate Testing Architecture

Figure 55 shows the proposed test architecture for a RF component embedded in a system with DSP resources. The resources depicted are common in highly integrated systems like SoCs and SOPs. Most of the time, these implementations already include support for boundary scan test of digital components. The proposed architecture makes use of this digital scan chain to capture the output of the comparator and to feed it into the available DSP resources. The algorithm depicted in Figure 54 can be implemented with these resources, which may be available in the form of FPGAs, ASICs or general purpose processors.

The additional BIST components consist of an analog multiplexer or switch, an oscillator with relaxed constraints on precision, a low-speed one-bit analog comparator, and a pseudo-random noise generator. The oscillator supplies the analog stimulus

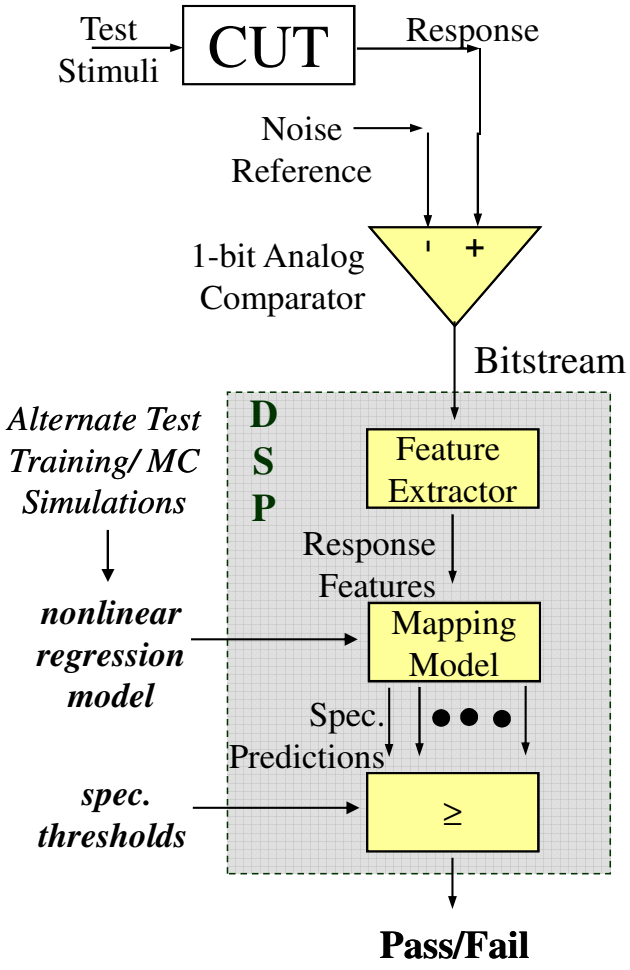


Figure 53: Noise-referenced alternate test setup.

for the DUT. It is tuned to a predetermined frequency and the linearity constraints are relaxed in a sense that the imperfections in the stimulus can be compensated by the methodology. Hence, it is not a critical design component and does not require a significant design effort if implemented in a DfT flow.

The pseudo-random noise generator makes use of a simple RC circuit and a LFSR to generate analog noise with specifications discussed in subsection 4.3.1. The analog noise is fed into the one-bit analog comparator, which operates at a frequency that is an order-of-magnitude below the necessary Nyquist rate. An averaging circuit can be employed instead of a sample-and-hold to relax the constraints on the design of

Notation:

x, xmc	: comparator output, digital bit stream
X, XMC	: reconstructed FFT
w	: vector of wavelet transform parameters
s	: vector of DUT specifications
nf_i	: FFT noise floor for i^{th} training instance
FE	: set of extracted frequency indexes
f	: vector of extracted features
M_j	: nonlinear regression model for j^{th} specification
t	: vector specification thresholds
D	: digital scan chain
p_i	: pass-fail decision for i^{th} specification
n_s	: number of specifications
n_{mc}	: number of Monte Carlo instances
P	: pass/fail decision for DUT

// generate feature extractors
$FE \Leftarrow U$
for $i = 1$ to n_{mc} **do**
$xmc_i \Leftarrow SimulateMonteCarloInstance(i)$
$XMC_i \Leftarrow ComputeFFT(xmc_i)$
$nf_i \Leftarrow ComputeFFTNoiseFloor(XMC_i, w)$
$FE \Leftarrow FE \cap j, for\ j = Index(XMC_i > nf_i)$
end for
// algorithms for DUT
$x \Leftarrow CollectBitStreamfromScanChain(D)$
$X \Leftarrow ComputeFFT(x)$
$f \Leftarrow X_j, for\ j \in FE$
for $j = 1$ to n_s **do**
$s_j \Leftarrow ComputeMapping(M_j, f)$
$p_j \Leftarrow ApplyGo/No\text{-}GoThreshold(s_j, t_j)$
$P \Leftarrow P \,\&\, p_j$
end for
return *P*

Figure 54: Algorithm for pass/fail calculation using DSP resources with noise referenced alternate testing.

the comparator. It compares the analog noise with the response of the DUT, and generates a one-bit digital output. The single bits at the output of the comparator do not carry information one-by-one, but when the sampling is carried over many cycles the resultant bit stream has an approximate imprint of the spectral content of the DUT response. This process is analogous to sampling the DUT response at random intervals and then extracting stochastic parameters from those samples.

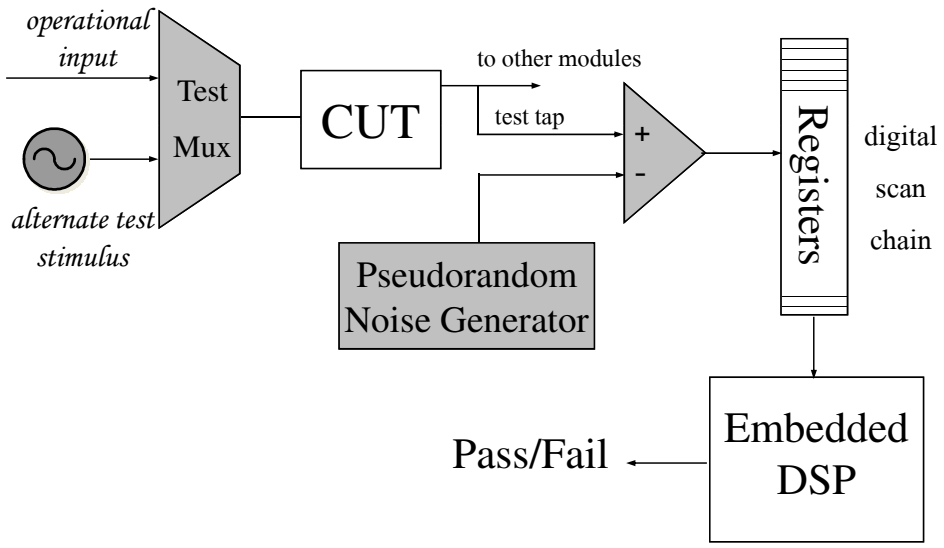


Figure 55: Implementation of noise-referenced alternate testing.

When the FFT is constructed at the DSP resources, the effect of random sampling dictates itself as a noise floor. Figure 56 shows the FFT semilog plot of a DUT response. Figure 57 shows the same response after noise-based comparison process. Higher order harmonics present in the original response are lost in the noise floor, while the fundamental harmonic is still visible with a skirt around it. In an ideal undersampling process, this skirt would not be present and the FFT would consist of a single amplitude value around each harmonic. In practice, the effects of accuracy problems in the sampling interval and in the sampling jitter dictate a skirt around the harmonics. From a different point of view, this skirt is a translation of the phase information in the transient signal to amplitude information in the FFT. Thanks to this translation process, alternate tests can use the amplitude information in the skirts to more accurately predict SUTs. The wavelet based automatic noise floor detection algorithm chops of the amplitudes below noise floor level, but the skirt amplitudes above this level are preserved and fed into the alternate test model for mapping into specification predictions.

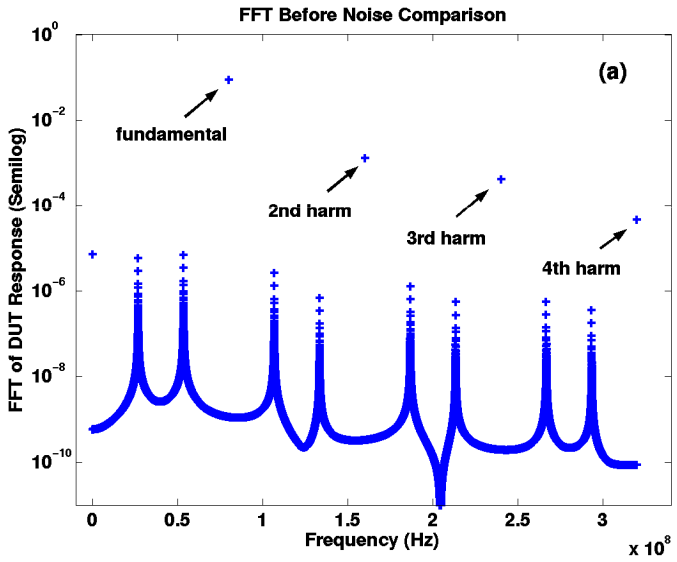


Figure 56: FFT of the output response before noise comparison.

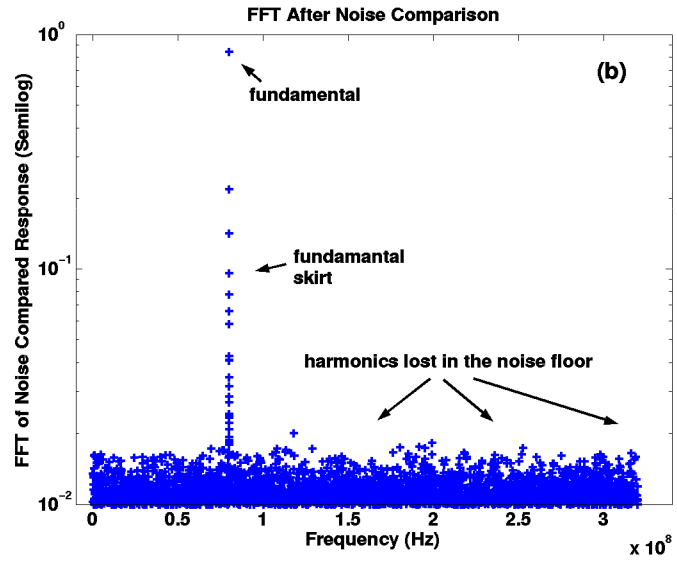


Figure 57: FFT of the output response after noise comparison.

A variation of this BIST architecture can be implemented as a BIT scheme for RF components without proper DSP support. The oscillator, pseudo-random noise generator, and the comparator are still embedded with the DUT, but the digital output of the comparator is latched and transferred to an external digital tester via a low-speed link. In this case, the feature extraction and mapping are implemented in

the tester. This variation proposes a low-cost alternative to high-speed RF testers.

4.3.3 Examples

Two simulation examples demonstrate the feasibility of the proposed approach. The bipolar LNA and mixer described in Chapter 3 are the DUTs.

For the LNA, the corresponding alternate test stimulus is a 900MHz single-tone sinusoid. The response is undersampled at 89.1MHz to generate 65536 samples of an effectively 7.2GHz signal, Nyquist rate of the 4^{th} harmonic. The higher-order harmonics are eliminated by the feature extraction algorithm, leaving only 10 samples of the skirt around the fundamental. The regression model is generated by 150 MC instances and specifications are validated using a separate set of 50 instances. Figures 58, 59, and 60 show the predicted versus actual specification values and prediction errors for IIP3, 1dBComp and NF respectively for these validation instances.

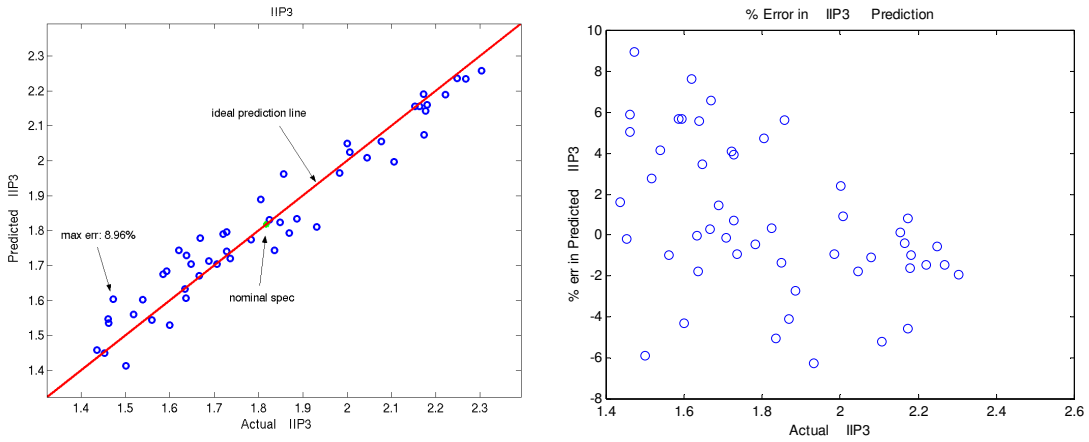


Figure 58: Predicted IIP3 and prediction errors for the LNA.

Table 4.3.3 lists the maximum prediction errors for this experiment in row 2, and errors from an ideal sampler in row 1. Ideal sampler is assumed to be a 12-bit ADC sampling at 7.2GHz. The maximum prediction error is 8.96% for the IIP3 specification, 2.38% for 1dBComp, and 0.799% for NF. Although these raw numbers are

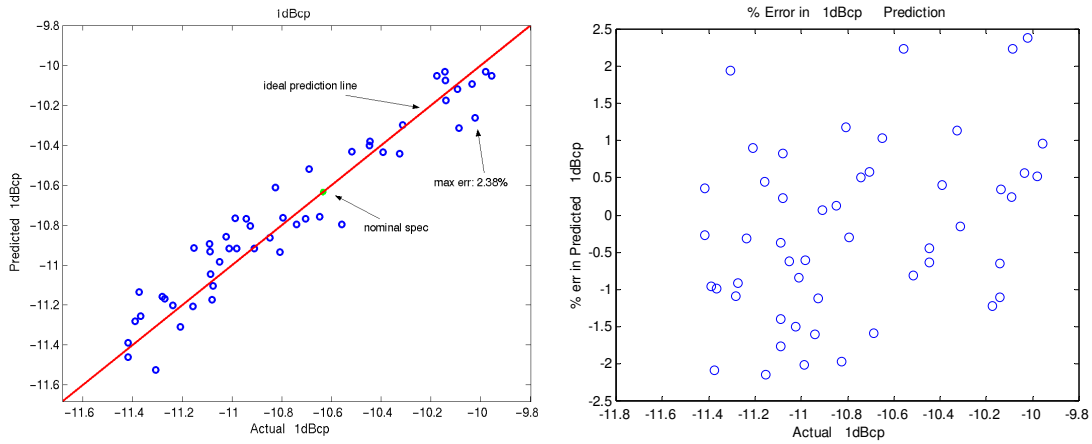


Figure 59: Predicted 1dBComp and prediction errors for the LNA.

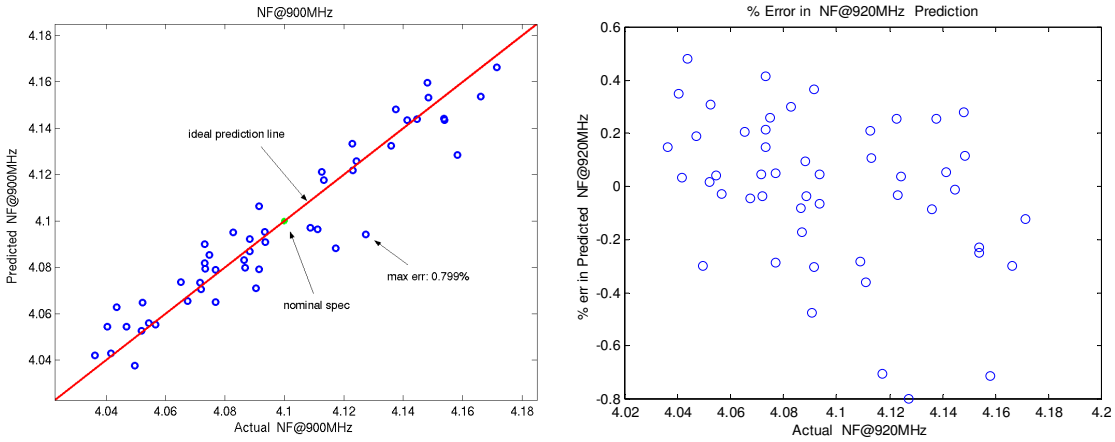


Figure 60: Predicted NF and prediction errors for the LNA.

significantly larger than those of the ideal sampler experiment, the ISO graphs show tight fits around the ideal 45° line. To see the impact of this increase in prediction error to classification accuracy, arbitrary pass/fail boundaries are introduced on a 10-step uniform grid along minimum and maximum ranges of each specification. Out of 500 cases there are only two misclassifications, both are normally passing instances classified as failing.

For the mixer, the corresponding alternate test stimulus is a 920MHz sinusoid accompanied by a 1GHz LO signal. The response, which has a fundamental at 80MHz,

Table 9: LNA maximum prediction errors in percentages of the actual specification values.

		IIP3	1dbC	Noise Figure
1	LNA Ideal	1.1%	1.02%	0.0190%
2	LNA BIST	8.96%	2.38%	0.799%

is undersampled at 71.1MHz to generate 65536 samples of an effectively 640MHz signal, which is the Nyquist rate of the 4th harmonic. The feature extraction algorithm removes the spectrum of the 2^{nd}, 3^{rd}, and 4^{th} order harmonics since they are below the noise floor. As a result, the only extracted feature is eleven samples of the skirt around the fundamental harmonic. These samples are fed into a regression model generated by 150 MC instances of the mixer netlist. The accuracy of the final test plan is validated by a separate set of MC instances that are generated independently from the training set. Figure 61 shows the predicted versus actual specifications of 50 MC instances in this validation set, and Figure 62 shows the associated prediction errors.

Table 4.3.3 lists the maximum prediction errors for this experiment in row two, and errors from an ideal sampler in row one. Ideal sampler is assumed to be a 12-bit ADC sampling at 640MHz. The maximum prediction error is 2.9% for the IIP3 specification, 1.56% for 1dBComp, and 1.14% for NF. Although these raw numbers are larger than those of the ideal sampler experiment, the ISO graphs show tight fits around the ideal 45° line. To see the impact of this increase in prediction error to classification accuracy, arbitrary pass/fail boundaries are introduced on a 10-step uniform grid along minimum and maximum ranges of each specification. Out of 500 cases, all instances are correctly classified.

The mixer experiment is expanded to validate the compensation in the presence of imperfect stimulus generation. The ideal sinusoidal stimulus is modified in different

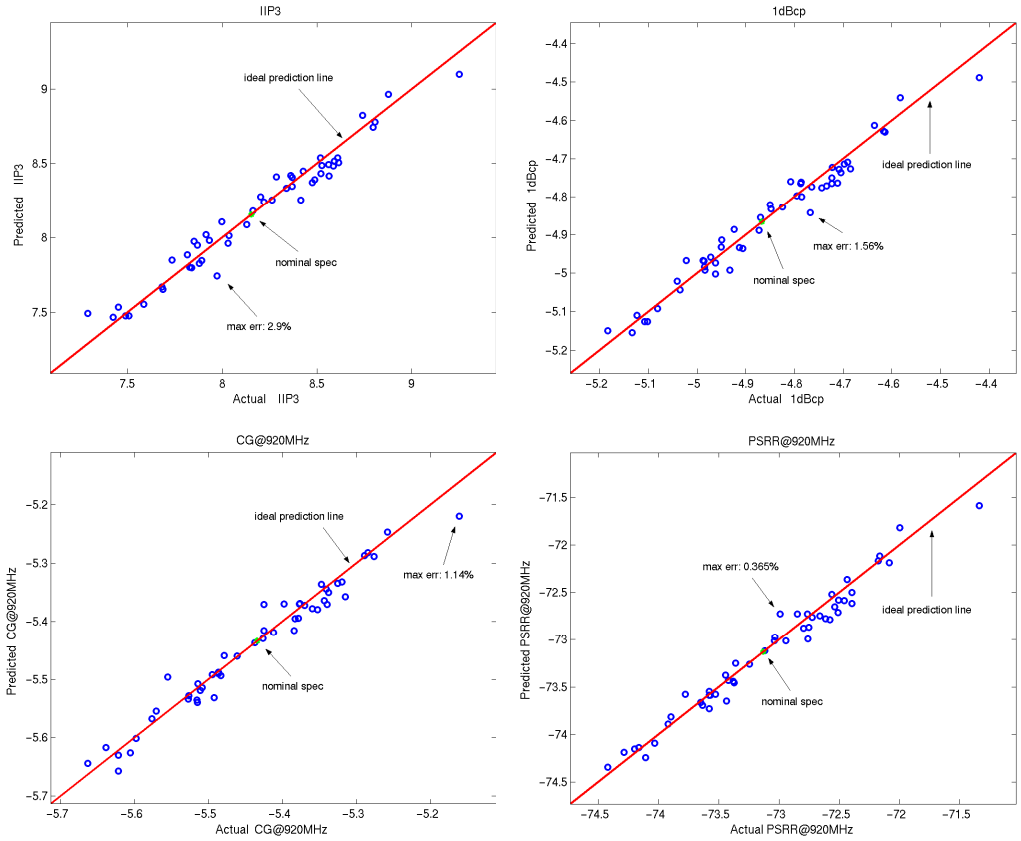


Figure 61: Predicted specification values for 1GHz downconverter mixer.

ways to emulate imperfections. The regression models and feature extractors generated by the ideal sinusoids are used for predicting specifications of the validation sets stimulated by imperfect sinusoids, no additional training instances are used.

- A positive DC bias is introduced to the stimulus with magnitudes ranging from 10mV to 200mV. As shown in Figure 63, noise-referenced alternate testing fully compensates for DC bias. No change in prediction error is observed.

Table 10: Mixer maximum prediction errors in percentages of the actual specification values.

		IIP3	1dbC	Gain	PSRR
1	Mixer Ideal	1.21%	1.02%	0.454%	0.112%
2	Mixer BIST	2.90%	1.56%	1.14%	0.365%

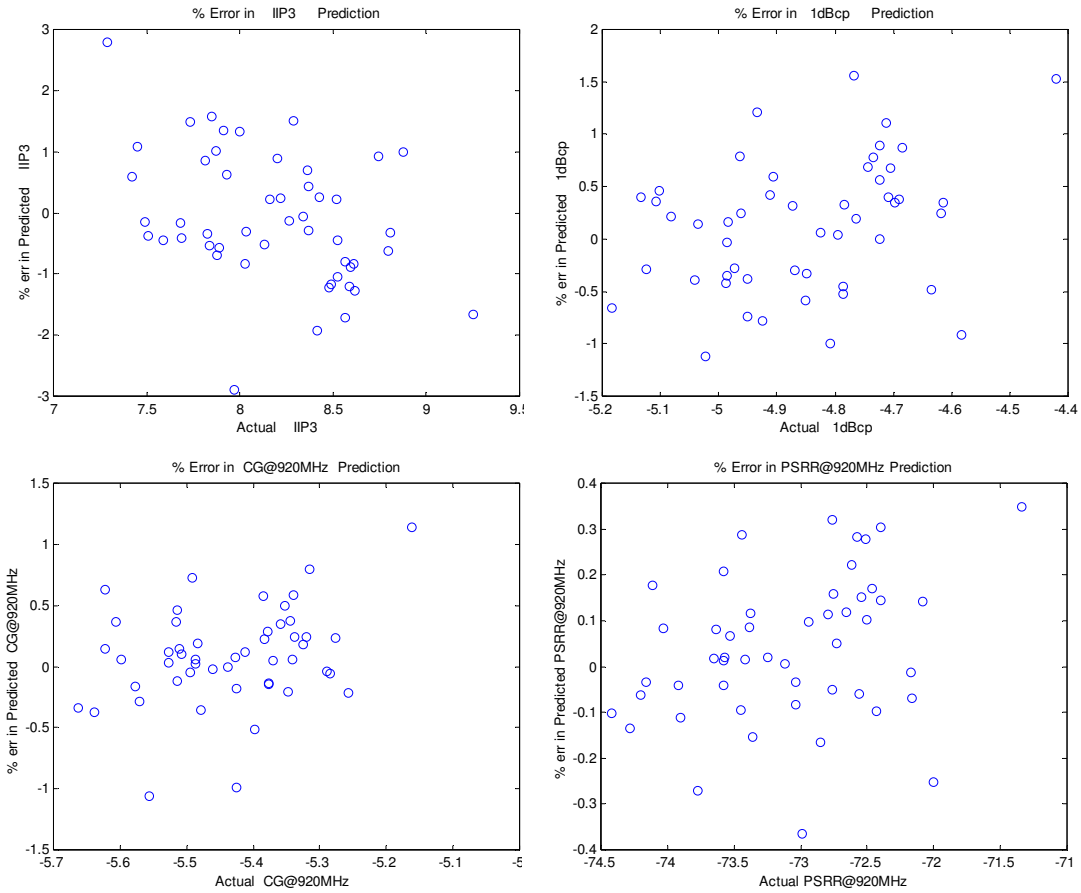


Figure 62: Prediction error values for 1GHz downconverter mixer.

- A positive phase shift is introduced to the stimulus with magnitudes ranging from 0.36° to 180°. Although the magnitude of the input spectral content does not change, the output spectral signature changes slightly because of sampling errors. Figure 64 shows the maximum error versus the amount of phase shift. The error introduced by phase shifts up to 3.6° is acceptable for many BIST applications.

- White noise is introduced to the stimulus with magnitudes ranging from 20mV to 200mV. Figure 65 shows the maximum error versus the amount of white noise. The error introduced by noise up to 40mV is acceptable for many BIST applications.

- A frequency shift is introduced to the stimulus with magnitudes 10kHz and 100kHz. Figure 66 shows the maximum error versus the amount of frequency shift. Although 10kHz difference does not show significant error difference, 100kHz difference introduced 23% error to IIP3. High sensitivity to frequency shift is expected since it changes the spectral content of the output, and moves skirt components significantly based on the interaction with undersampling scheme.

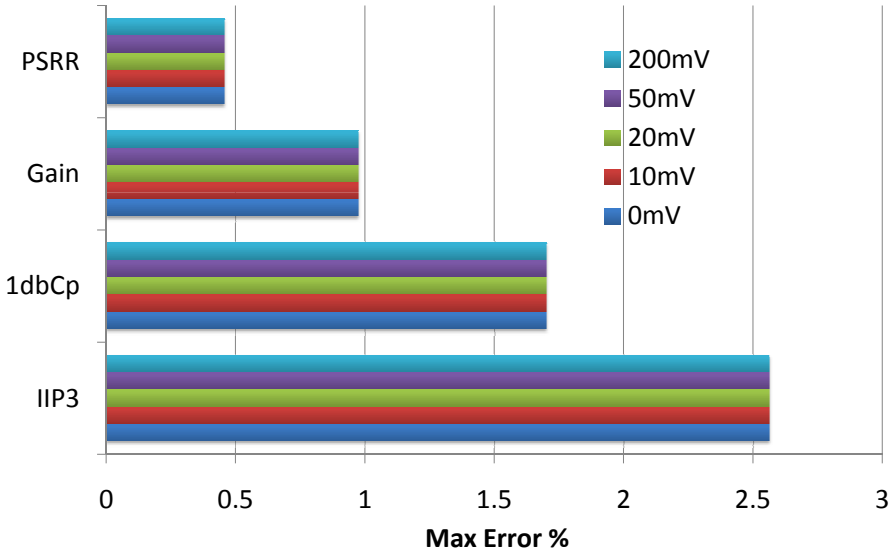


Figure 63: Prediction errors with DC shift in stimulus.

4.4 BIT of RF Components with Mapped Feature Extractors

In this section, specialized functions of the output response from an alternate test are computed using built-in feature extraction sensors, which measure a complex function of the response waveform and output a DC signature. Different sensor structures are

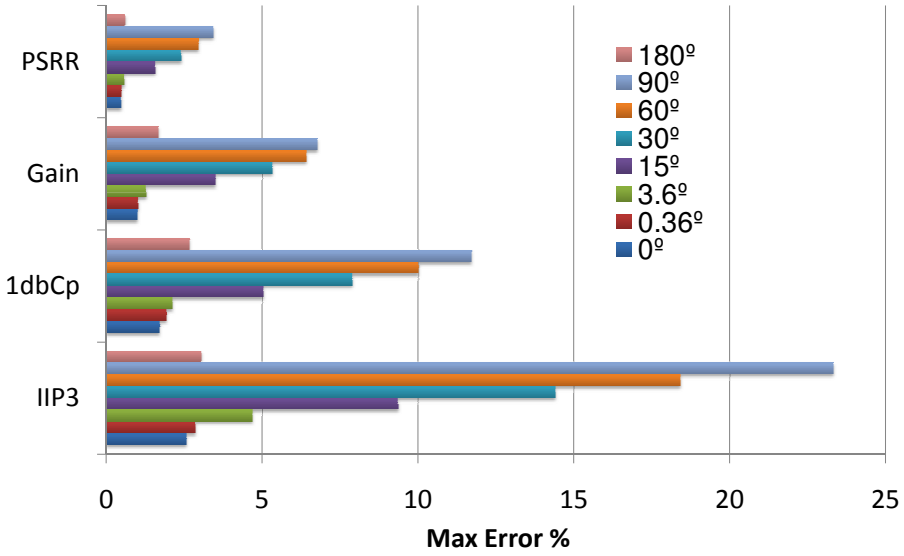


Figure 64: Prediction errors with phase shift in stimulus.

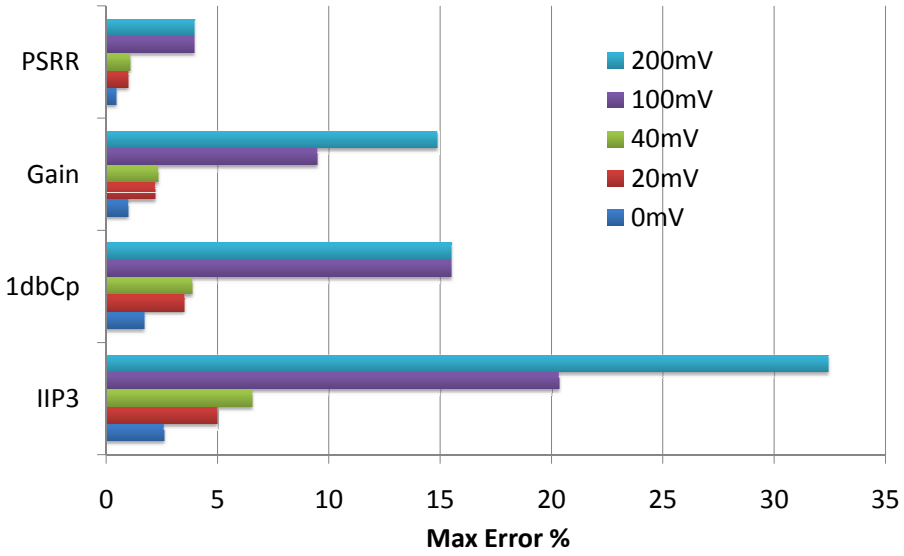


Figure 65: Prediction errors with white noise in stimulus.

evaluated based on their performance in the presence of environmental effects and process shifts. It is shown that very simple sensing circuitry can produce high quality alternate tests for RF components.

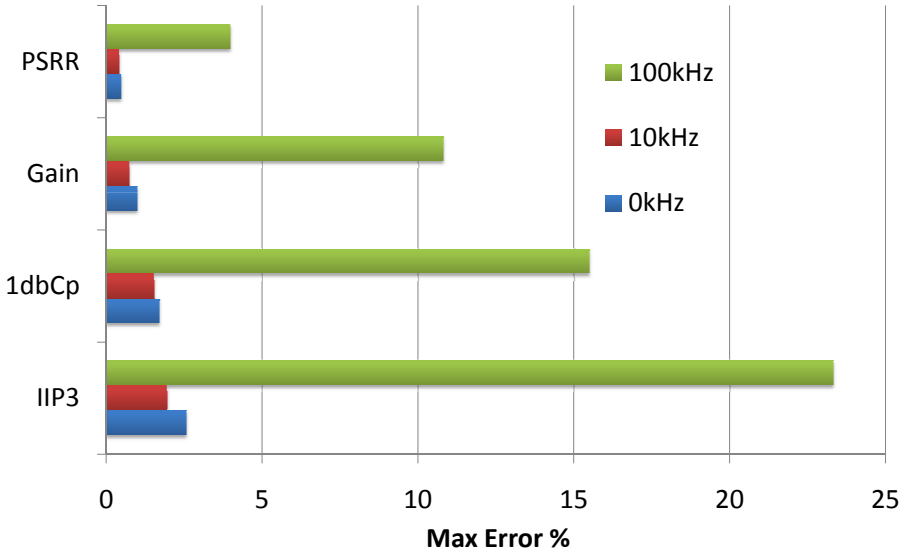


Figure 66: Prediction errors with frequency shift in stimulus.

DC-level feature extraction has been used in test setups to measure properties of electrical signals such as bias current/voltage, peak, rms, zero-crossing and tuned spectral components. Each feature extractor, or sensor, is built specifically for the single property it targets. A recent example in [129] demonstrates an approach, which can generate quantitative measures for more than one specification; however, the measurement-to-specification mappings have to be hand-crafted specifically to the characteristics of the DUT. The circuitry associated with such detectors is generally complex, and the DC voltage generated is only an approximation to the signal feature measured that is valid in only specific ranges. For example, an rms detector generates a DC voltage proportional to the rms value of the signal, but its accuracy is limited with assumptions such as periodicity, waveform shape, or maximum/minimum amplitude. This is generally because of the non-linearity present in the devices used to design the detector. Even bipolar-based applications require post-production calibration for accuracy. To make matters worse, on-chip feature extractors suffer from the same process variation and thermal effects that impact the DUT. Making the sensor

more complicated and robust usually means that the sensor size approaches the RF DUT size, shifting the test focus from the DUT to the detector.

Alternate testing methodology can provide a solution to these challenges. Specification-based alternate tests provide a general methodology independent of the DUT or the target specifications. In this sense, it is a complete tool, which can be applied across different devices, and multiple specifications can be verified through a single hardware minimizing area overhead. [130] proposed a detector-based implementation, which makes use of hardware-based test response feature extractors to produce a DC signature of the alternate response. Although this experiment demonstrates the potential of DC level sensors used together with alternate tests for BIT, explicit features, such as peak or rms, require the use of complex circuitry for measuring the corresponding peak and rms values.

To alleviate the above problems with conventional DC-level sensors, such as rms/peak detectors, we propose a new class of sensors, *mapped feature extractors*, to be used with alternate test methodology. There are two key contributions of this work:

- Simple sensor structures instead of conventional rms or peak detectors. The conventional detectors are very accurate but complex in structure hence occupy large area, comes with significant parasitics, and need elaborate calibration schemes to satisfy the region of operation. In contrast, mapped feature extractors are simple, and yet able to yield accurate specification prediction with alternate tests.

- A methodology that can compensate these simple on-chip sensors for environmental effects. Examples include temperature and process variations compensation without adding complexity to the feature extractor.

4.4.1 DC-Level Feature Extractors

DC-level feature extraction is not a new concept; the idea has been used in measurement setups in order to study the properties of a waveform, which is otherwise not completely characterized. The most well known examples of such extraction circuitry are in the form of peak or rms detectors. Other common examples are zero crossing detectors, bias current/voltage sensors and tuned spectral component detectors. These extractors are common because the DC values they represent can be easily related to a physical property of the original AC waveform. In other words, they provide an explicit relationship between the detector output and the DUT response. Such DC properties and their corresponding sensor structures, explicit feature extractors, can be directly used with alternate test to produce predictions for SUTs [130].

There are many challenges in the implementation of explicit feature extractors. The DC values they present are almost always approximations to the original feature under a long list of assumptions such as waveform type, frequency and swing ranges or piecewise fitting. The deviations are usually a result of the nonlinearity present in the semiconductors making up the detector. In this sense, highly linear devices such as bipolar transistors are preferred over field effect transistors, which generates a bottleneck given that the range of bipolar devices available in modern processes are very limited. Even bipolar based applications require a calibration scheme in post-production, when the accuracy is a key element. This calibration step is reflected to the manufacturing cost of the device in terms of extra time and real estate overhead. Furthermore, when these sensors are built-in, they are affected by the same process variations and environmental effects that degrade the performance of the DUT. In order to make them immune to these variations, the designer can add additional circuitry; however this approach makes the sensor larger and more complicated. If the DUT is an RF component, where simplicity is a key element, the sensor may become even larger than the device, shifting the focus to the sensor rather than the

DUT since the former will be more susceptible to faults.

As an example, let's consider the peak detector proposed in [131], as shown in Figure 67. This peak detector is made up of two bipolar transistors, which are arranged to produce a differential output, one side of which is used to cancel the DC bias. This implementation is selected as an example because its differential nature helps protect against process and environmental variables; furthermore, it provides a linear mapping when compared to field-effect transistor based or diode-based peak detectors, yet keeping a simple structure with low transistor count. Following all these properties, it represents a high-end example for common explicit feature extractors. However, even this implementation needs calibration for specific regions of operation. Moreover, its output is proportional to the peak of the signal provided that the signal is a sinusoidal. For distorted waveforms its accuracy fades dramatically destroying the one-to-one mapping. In [132], the accuracy of this circuit is discussed, and it is shown that the relative error in representing the peak value goes up to 900% for a transition region in its transfer characteristics, as shown in Figure 68. So, any application making use of this detector should implement a calibration procedure for this region. Subsequenctly, a hardware modification is proposed to implement a more linear transfer function, keeping its relative error to 8%. On the other hand, this modification makes the implementation more complicated, and adds significant area overhead. Our experiments suggest that in the presence of regular process and temperature variations, the relative error for the original circuit is 42% excluding the problematic transition region, and the error goes up to 63% after hardware modifications proposed in [132]. Hence, the performance of this detector in a built-in test environment is unsatisfactory for the common process and environmental variations affecting the original DUT.

The problems depicted for the example above are characteristic for explicit feature

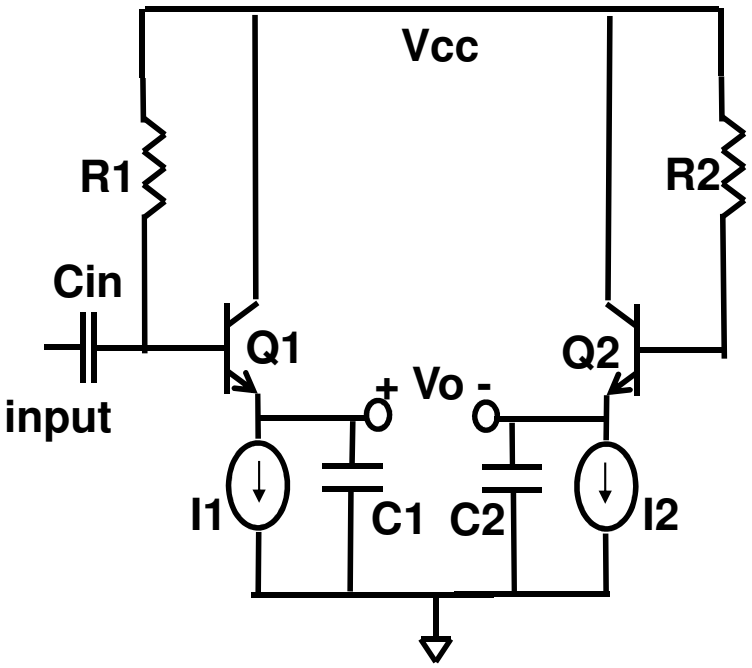


Figure 67: Bipolar peak detector in Meyer.

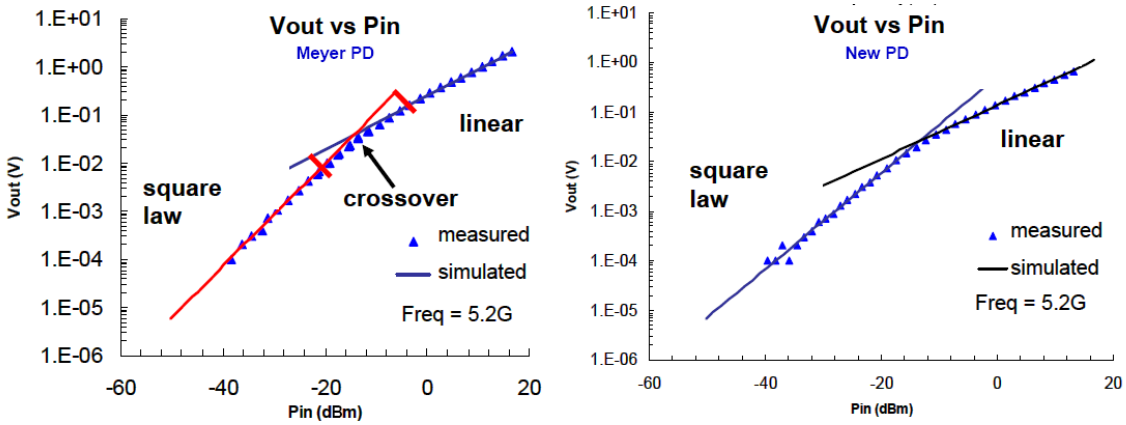


Figure 68: Error in the bipolar peak detector of Meyer and correction proposed in Zhang et.al.

extractors. In this case, the proposed sensor functions as a peak detector only under a long list of assumptions, which can easily be violated in a BIT environment even when some form of calibration is provided. As a matter of fact, the DC signature for this sensor presents a richer content than a single peak value. In favor of being able to build the correlation relation easily, the user constrains its potential ability by settling

down to an approximation of a one-to-one mapping by ignoring its nonlinearity and variation under non-ideal sinusoidal inputs. In contrast, if the output of such a sensor is used with alternate test methodology, the inherent mapping process will make better use of its richer signature content instead of treating it as a distorted approximation of a peak value. In other words, the implicit features of the signature are made explicit by the mapping process. Moreover, the process provides inherent calibration for problematic operating regions as well as process and environmental variations. By using alternate tests with very simple sensor structures, one can have high-end results, which are otherwise only possible with elaborate post-production calibration and additional complex circuitry to take care of variation effects.

This approach presents a new sensor paradigm for use with alternate test methodology. Since the measurements in alternate tests are different from those made in classical specification based tests, the built-in sensors used for measuring classical figures of merits such as peaks, root-mean-square values, zero-crossings, etc can be replaced with implicit feature extractor hardware that measure figures that are more accessible but harder to relate to the specification value. Although this relation is not easy to build, the mapping process in alternate test will build a good enough prediction as long as the changes in the measured figure are correlated to the SUTs. In order to validate the idea, we take the simple peak detector in [131], and study its signature characteristics when its input is not bounded by assumptions. After studying these implicit features, the sensor is further modified to generate a class of sensors, which can be used together for predicting complex device specifications such as IIP3, 1dBComp, and NF instead of a single peak value. Furthermore, we show that alternate test provides an inherent calibration when the sensors as well as the DUT are subject to large temperature variations.

4.4.2 Differential-Topology Sensor Class

Figure 67 shows the peak detector in [131], where $R1 = R2$, $C1 = C2$, and $I1 = I2$. The input waveform can be dissected into a bias voltage and $x(t)$ such that

$$V_i = V_B + x(t), \tag{12}$$

and

$$V_C = V_B - V_A; \tag{13}$$

$$V'_C = V_B - V'_A. \tag{14}$$

Then for Q1 and Q2 being identical:

$$I_{C1} = I_{S1} \cdot e^{\frac{V_C}{V_t}} \cdot e^{\frac{x(t)}{V_t}}, \tag{15}$$

$$I_1 = I_{S2} \cdot e^{\frac{V'_C}{V_t}}, \tag{16}$$

$$I_{C1}\Big|_{DC} = I_{S1} \cdot e^{\frac{V_C}{V_t}} \cdot e^{\frac{x(t)}{V_t}}\Big|_{DC} = I_1, \tag{17}$$

$$e^{\frac{V_C}{V_t}} \cdot e^{\frac{x(t)}{V_t}}\Big|_{DC} = e^{\frac{V'_C}{V_t}}, \tag{18}$$

$$e^{\frac{V'_C - V_C}{V_t}} = e^{\frac{x(t)}{V_t}}\Big|_{DC}. \tag{19}$$

Since $V_o = V_A - V'_A = V'_C - V_C$,

$$V_o = V_t \cdot ln\left(e^{\frac{x(t)}{V_t}}\Big|_{DC}\right). \tag{20}$$

The derivations in [131] and [132] depend on the assumption that when $x(t)$ can be represented with a sinusoidal, a modified Bessel function can be used to compute an approximation of the DC value for $e^{\frac{x(t)}{V_t}}$. Equation 20 is a generalized version of this derivation without any additional assumptions.

Although the detector in Figure 67 can be used with alternate tests to predict the peak value of the sensor input, prediction of more complex specifications such as IIP3 or NF demand extra dimensions for the measurement space. Instead of carefully searching for another detector, we use a generic way to generate a class of sensors from a single architecture. Equation 20 is in the form of

$$V_{o1} = f_1(mean(e^{g_1(input)})), \tag{21}$$

where exponential characteristics come from the bipolar transistor. Figure 69 shows the second detector, when the bipolar devices are replaced with FETs. These two sensors, differential-topology sensors, make use of the same topology but with different active components. In this case, the logarithmic/exponential relation given in Equation 21 is replaced with a square-root/square relation yielding

$$V_{o2} = f_2(mean(g_2(input)^2)). \tag{22}$$

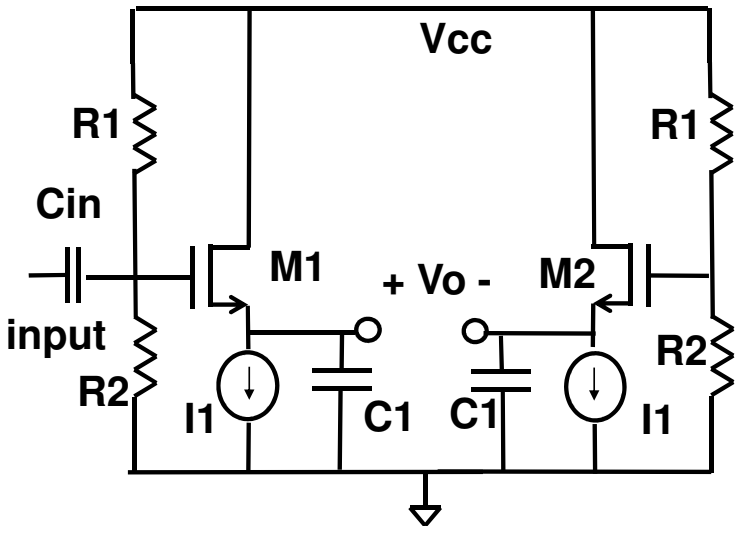


Figure 69: FET version of "peak detector".

Figure 70 shows a generic BIT setup using these sensors: a simple on-chip or on-package analog signal generator applies the test stimulus to either the DUT or the input sensors through a test multiplexer; the embedded output sensors together

with optional input sensors produce DC values to be sampled by the low-cost external tester. These DC values are fed into the specification mapping module in the external tester and non-linear mapping functions output predictions for SUTs.

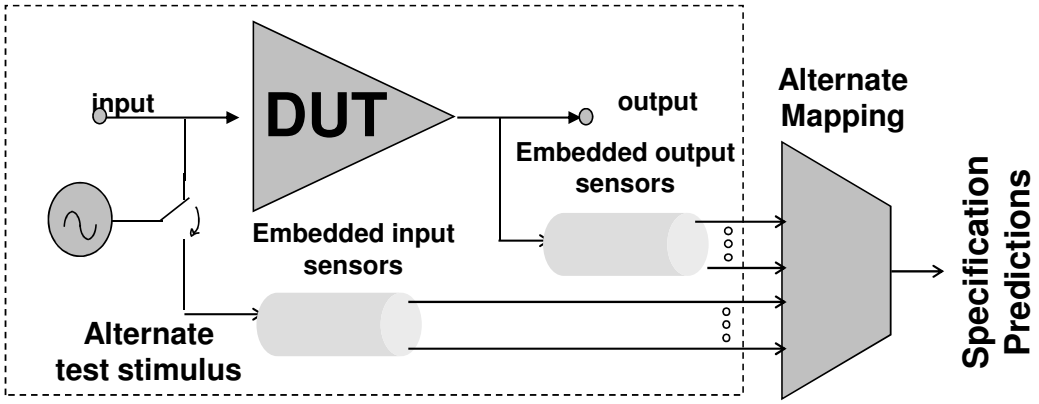


Figure 70: BIT setup with mapped feature extractors at the input and output.

The accuracy of the differential-topology sensor architecture is demonstrated by a series of simulation experiments using the 900MHz LNA described in previous sections. Five process variables with $3\sigma = nom/10$, where nom represents the nominal value for the variable, and σ is the standard deviation, are considered for the MC experiment. The process variables are saturation currents and the forward gains of the transistors, and the sheet resistance. The sample specifications of interest are 1dBComp, IIP3, and NF at the nominal operating frequency and temperature. The corresponding alternate test stimulus is selected as a single 900MHz sinusoid in favor of its simplicity to be generated on-chip/package by an LO or be supplied from a low-end external source.

Two sets of device instances are generated for training and validation purposes using the circuit netlist, device models, and process variable distributions. SpectreRF simulator is used to simulate all of these instances at the nominal operating frequency and at the nominal temperature of operation. These simulations are designed to measure actual specifications of interest for each circuit instance by classical

methods. Using the measurements and specifications from the training set, a set of nonlinear mappings are generated using MARS. Then, these mapping are used with the validation set to check the accuracy of predictions for each specification under test.

The experiment is designed to be performed in six steps, each investigating a controlled branch in the space of possible experiments. The end goal is to demonstrate the auto-calibration ability of the methodology in the presence of temperature variations by using the two differential-topology sensors at the output of the DUT together with a third sensor at the input. First step checks prediction errors for the LNA, when the DUT analog response samples are used directly to generate the regression models and to predict the specification values of the validation set instead of differential-topology sensor outputs. Although sampling at that frequency is not feasible for a BIT application, these results represent an ideal limit for alternate test predictions without DC-level feature extractors, and are listed for comparison. Similarly, second step uses analog response samples only this time for the validation of the auto-calibration ability. For every auto-calibration experiment, the 100-instance training set is simulated at six discrete temperature values -20, 0, 20, 27, 40 and 60 °C; then, a new 400-instance validation set is generated by four copies of the original 100-instance validation set. Each instance in this new validation set is simulated at a random temperature in the range [-20 °C 60 °C]. The flow of the experiment is outlined in Figure 71.

Third step of the experiment implements a variation of Figure 70, with the two differential-topology sensors connected only at the output and no input sensors, as shown in Figure 72; while fourth step runs the temperature auto-calibration experiment described above with the same setup. The fifth step challenges the ability of the FET-based sensor as an explicit temperature monitor; in this experiment, the simulation temperature is provided to the training set explicitly, and MARS mappings are

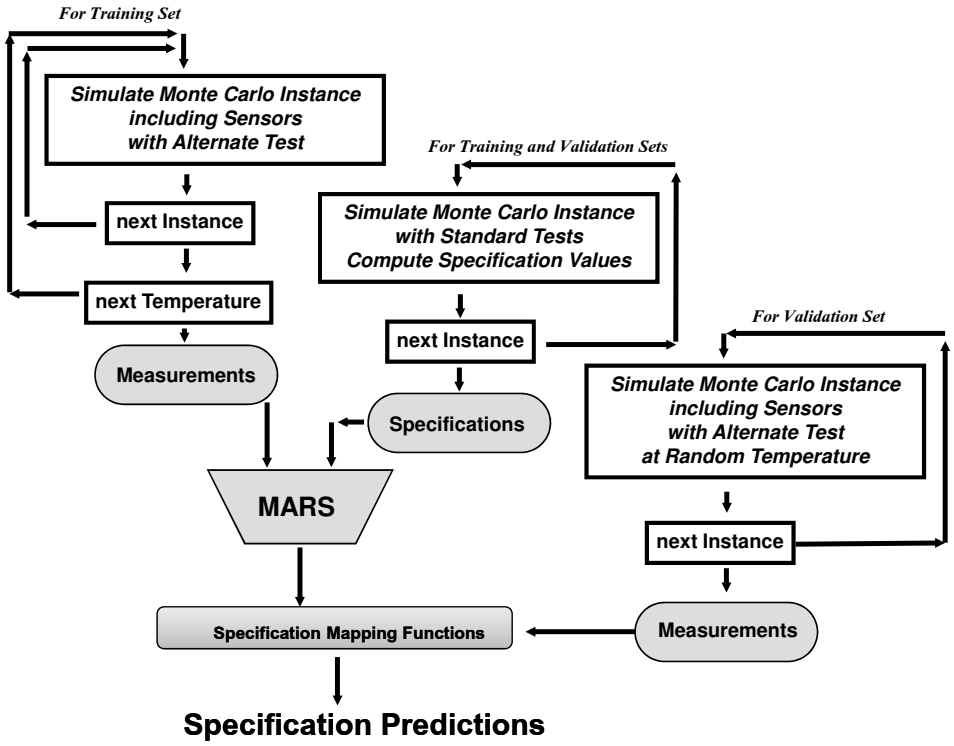


Figure 71: Algorithm for model generation and application.

generated for the temperature using this sensor alone, as shown in Figure 73. Finally, the sixth step of the experiment validates the proposed auto-calibration methodology by using the FET-based sensor at the input together with two differential topology sensors at the output.

Table 4.4.2 shows the summary of results for all six steps. For each case, the maximum prediction error is listed as the absolute difference from the original specification. The numbers in Table 4.4.2 should always be considered together with secondary measures such as percentage errors and number of misclassifications.

In order to validate the ability of differential topology sensors to predict complex specifications, one can compare the results of steps one and three. Both of these experiments are performed at a constant temperature. In the ideal sampled case of

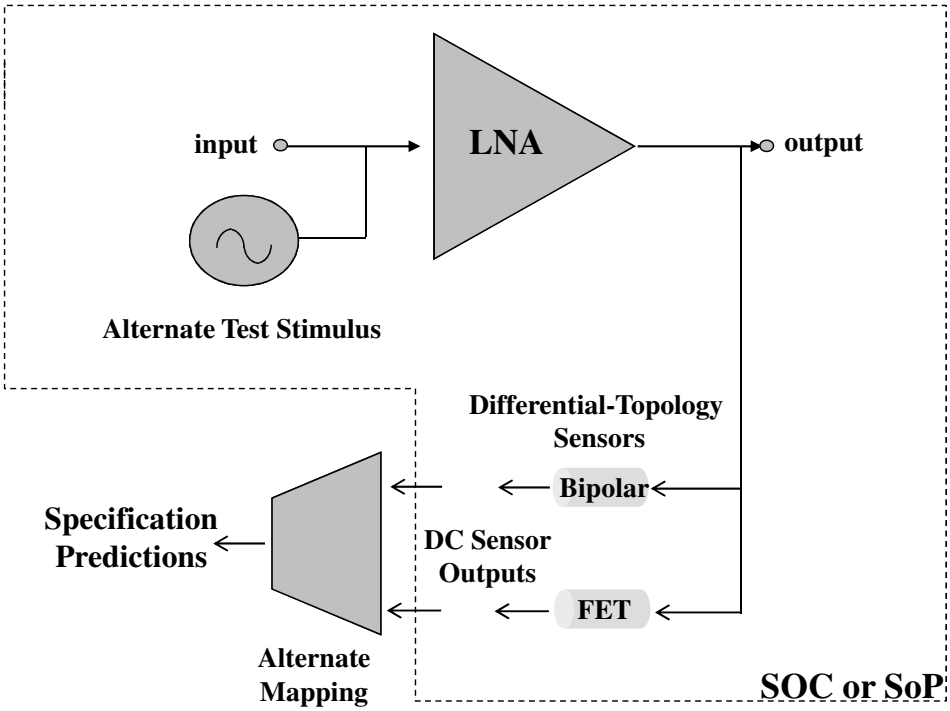


Figure 72: BIT setup with mapped feature extractors only at the input.

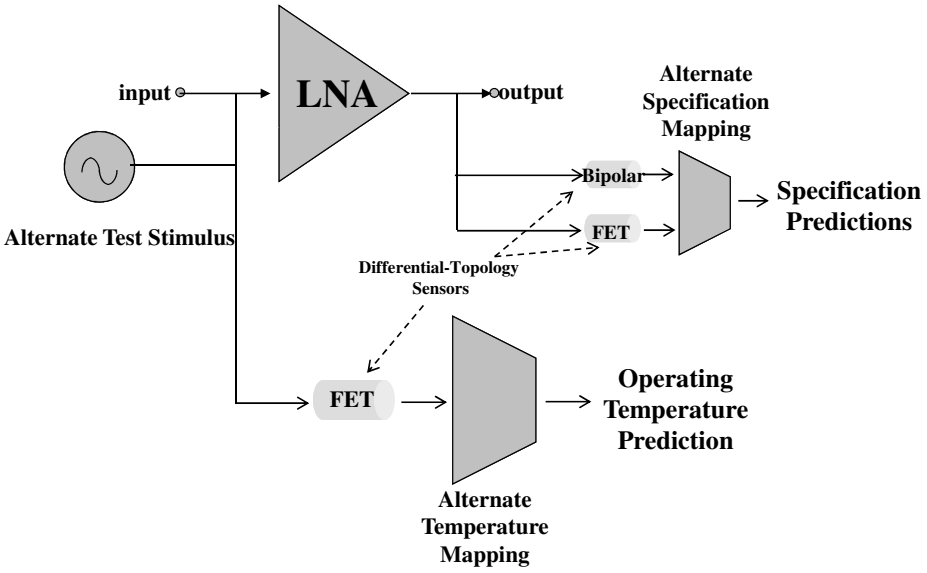


Figure 73: BIT setup for explicit temperature prediction.

step one, the maximum percentage error is 1.1%; whereas in step three using DC signatures of the differential topology sensors, the error goes up to 6.2%. Although this error is significantly larger than the ideal one, the accuracy is still comparable to the

Table 11: Maximum prediction errors of the actual specification values

#	Temp		IIP3	1dbC	NF
1	No	Ideal sampling	0.072 dB	0.092 dB	0.0081 dB
2	Yes	Ideal sampling	3.2 dB	4.8 dB	1.19 dB
3	No	2 sensors	0.41 dB	0.52 dB	0.25 dB
4	Yes	2 sensors	3.3 dB	4.7 dB	1.23 dB
5	Yes	2+1 sensors	3.3 dB	4.7 dB	1.23 dB
6	Yes	2+1 sensors	0.62 dB	0.94 dB	0.19 dB

error resulting from the repeatability of a classical test measurement. Furthermore, the misclassification rate is the same for both setups, only 1 out of 100 instances. The ISO graphs are shown in Figures 74, 75, and 76.

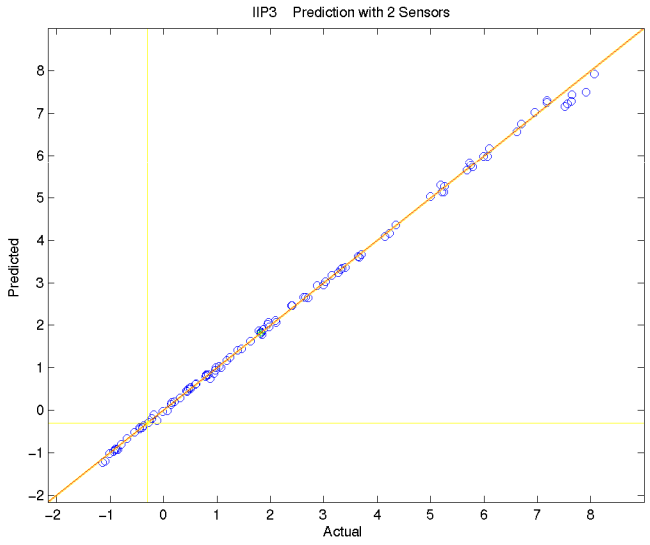


Figure 74: Predicted vs actual specifications for LNA IIP3 with two differential topology sensors at the output, no temperature variation.

When setups missing the temperature monitor sensor are compared with the corresponding setups performed at constant temperature - step two vs one, and four vs three -, the error percentages are observed to go up significantly, yielding similar misclassification rates around 21%. Although the predictions for the instances simulated

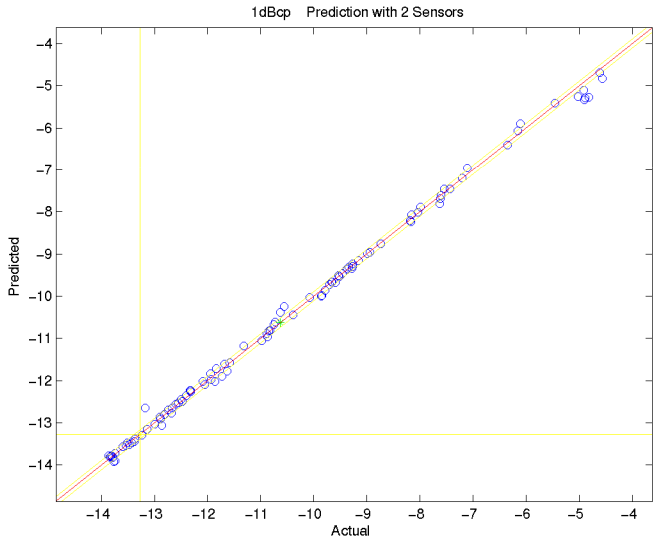


Figure 75: Predicted vs actual specifications for LNA 1dBComp with two differential topology sensors at the output, no temperature variation.

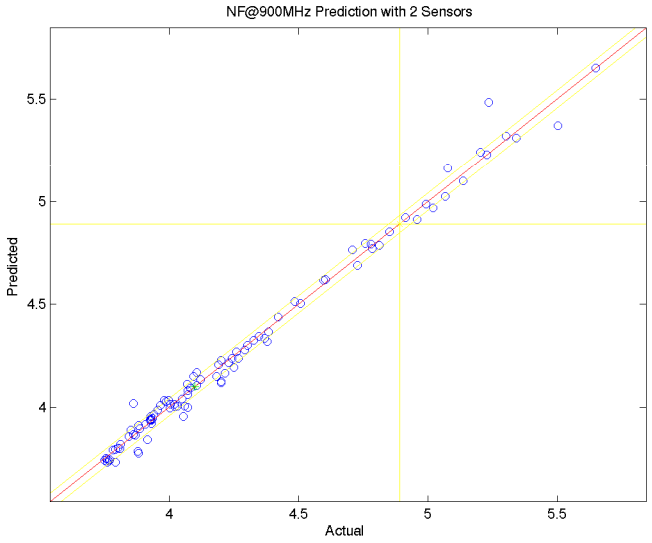


Figure 76: Predicted vs actual specifications for LNA NF with two differential topology sensors at the output, no temperature variation.

at around the nominal temperature are similar in terms of accuracy, the rest of them result in significant deviations from the 45° line. This hazy constellation graph is depicted in Figure 77 for IIP3 measurements in step four. The figure shows that in the absence of the input sensor acting as a temperature monitor, the errors are large

as expected.

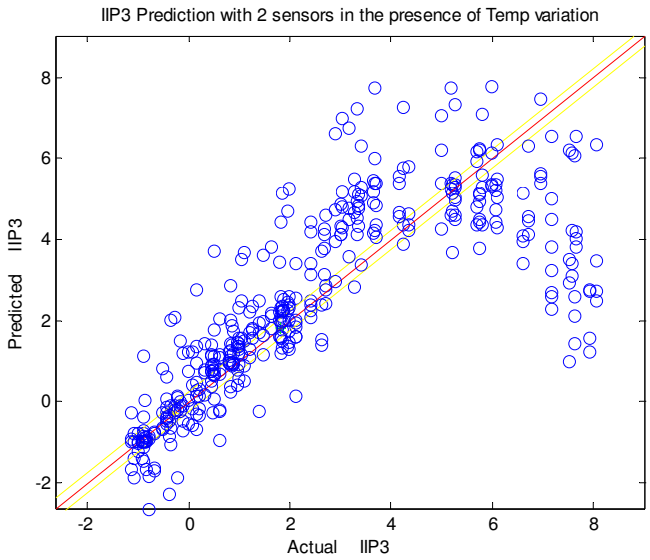


Figure 77: IIP3 prediction without temperature monitor.

Steps five and six are performed in the presence of an input sensor as a temperature monitor. In step five, the signature of this additional sensor is only used for prediction of temperature as an explicit goal; hence, the specification predictions are not different from those in step four. The purpose of step five is not to enhance the specification prediction, but to validate the use of the additional sensor as a temperature predictor. The results from the temperature mapping module show that the maximum error is 3.37 °C and the rms error is 1.20 °C, as shown in Figure 78.

Finally, step six validates the proposed auto-calibration methodology. In this case, temperature is treated as an internal variable, and the DC signature of the third (input) sensor is used with the other two output sensors to directly predict SUTs. All three readings are directed to the same mapping function trained to predict IIP3, 1dbComp and NF. Figures 79, 80, and 81 shows the ISO graphs for this setup, where the maximum percentage error is 8.1% and only 3 instances are misclassified out of

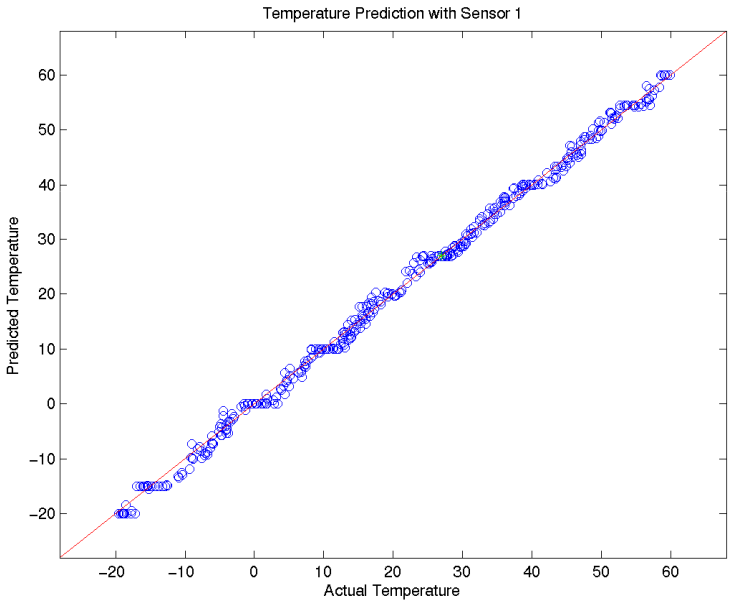


Figure 78: Predicted vs actual temperature with FET-based sensor used as a temperature monitor.

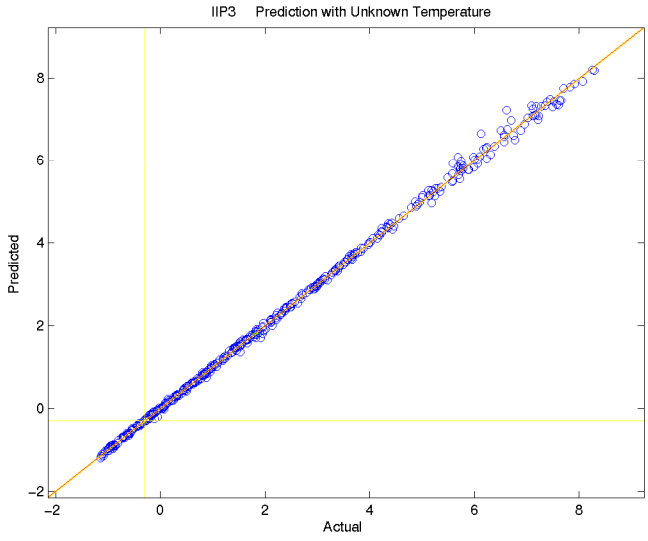


Figure 79: Predicted vs actual specifications for LNA IIP3 with two differential topology sensors and a temperature monitor sensor.

400.

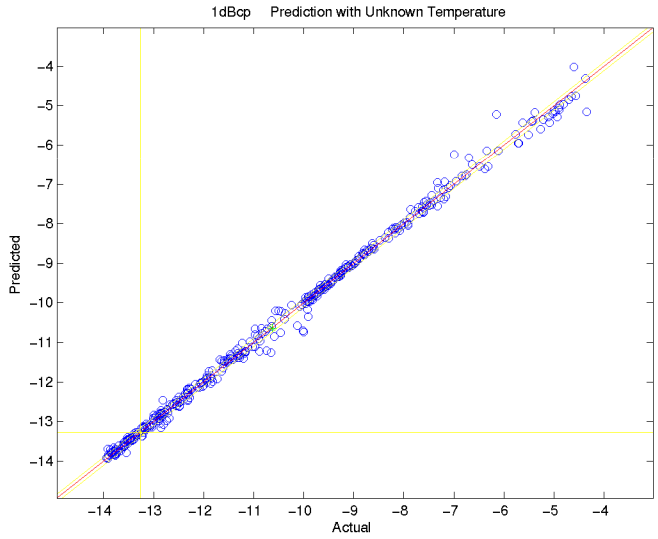


Figure 80: Predicted vs actual specifications for LNA 1dBComp with two differential topology sensors and a temperature monitor sensor.

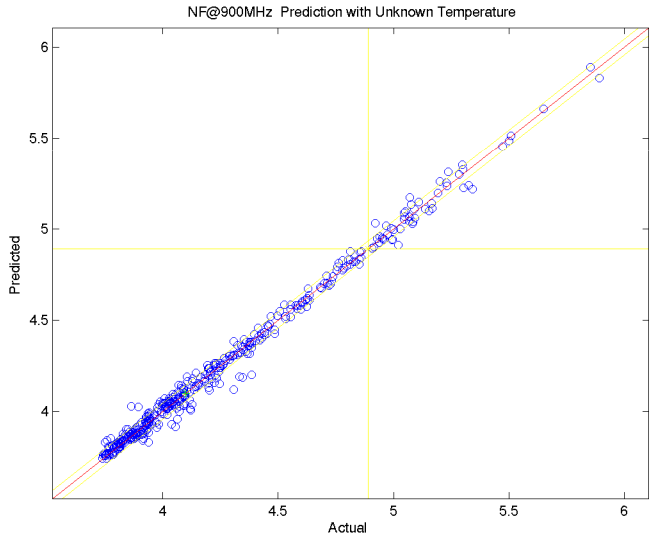


Figure 81: Predicted vs actual specifications for LNA NF with two differential topology sensors and a temperature monitor sensor.

4.4.3 Recursive Sensors

The mappings generated by alternate tests are valid under the assumption that the process variations on the manufacturing line are approximately within the same range used for producing the training set. When there is a large shift in one of the process

variables, the mappings need to be calibrated accordingly. Although a larger training set can be used to account for these process shifts, the number of instances in the training set grows exponentially with the range of variations. Instead, the input sensor concept as explored in the previous subsection can be used as process monitors, which can map process shifts efficiently with a small number of training instances.

To build process monitors, we first start with the differential topology sensors and evaluate their potential. Since these sensors are differential, their sensitivity is limited in terms of reflecting large process variations. Each differential topology sensor makes use of a different active device. This is a disadvantage for some processes, where only one type of an active device is present. Furthermore, each sensor at the output changes the load of the DUT and lumped implementations may require a redesign of the matching network.

Figure 82 shows a variation of the sensor in Figure 67. This new class of detectors is single-ended and the loading is constant regardless of the number of sensors at the output, because each new sensor input is connected to a node of the previous sensor. In this sense, the output function of sensor m, is defined recursively in terms of the sensors 1 to $m-1$ connected between the DUT output and input of the last stage. Furthermore, these sensors require only one type of active device. In this section, we will call n such sensors as an n^{th}-order *process monitor* if they are connected at the input of the DUT, and as an n^{th}-order *recursive sensor* if connected at the output. The question of using what combination of process monitors and recursive sensors is a debate between the improvement in accuracy and the extra area overhead, being subject to the magnitude of anticipated process shifts. Although the signatures of different orders are not independent, a certain level of redundancy helps avoid over-fitting to noisy signatures.

In order to compare the performance, we ran the same experiment in Table 4.4.2, step number three by replacing the two differential-topology output sensors with a

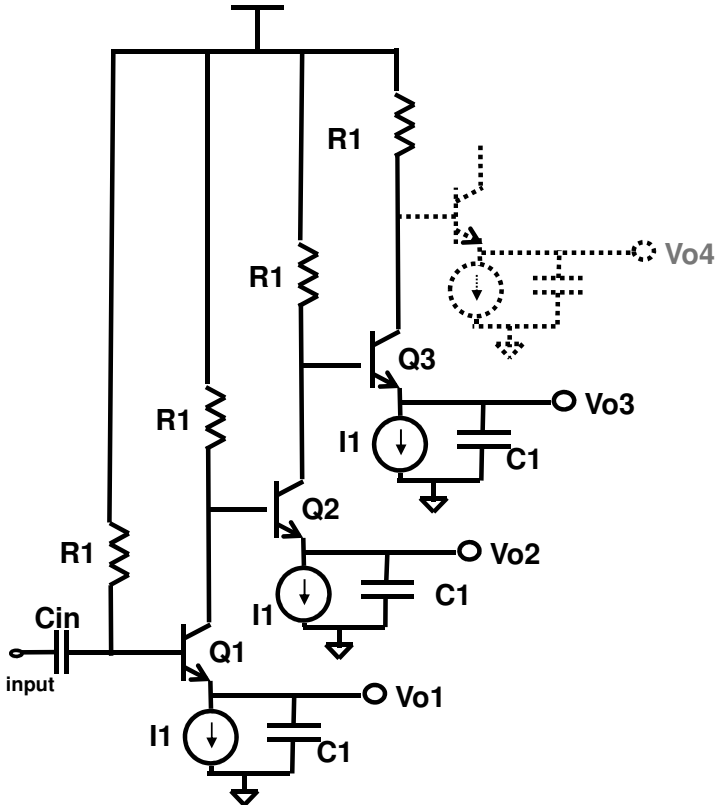


Figure 82: Recursive sensor.

2^{nd}-order recursive sensor. The maximum errors are very similar: 0.46dB, 0.54dB and 0.25 for IIP3, 1dBComp and NF respectively.

The process shift calibration capability of recursive sensors is demonstrated by a series of simulation experiments using the same LNA. In this case, each process variable is assumed to have a normal distribution with $3\sigma = nom/2$, instead of the $nom/10$ span. With this change, the process parameter space is enlarged by $5^5 = 3125$ times, which would require the same order of increase in training set if there were no process monitors. We have designed the experiments by using a 2^{nd}-order process monitor together with a 2^{nd}-order recursive sensor,as shown in Figure 83. Although, we have experimented with many other combinations and orders, the accuracy does not increase significantly above the $2x2$ configuration. To compare the results, we have also run experiments with a 2^{nd}-order recursive sensor and no process monitor.

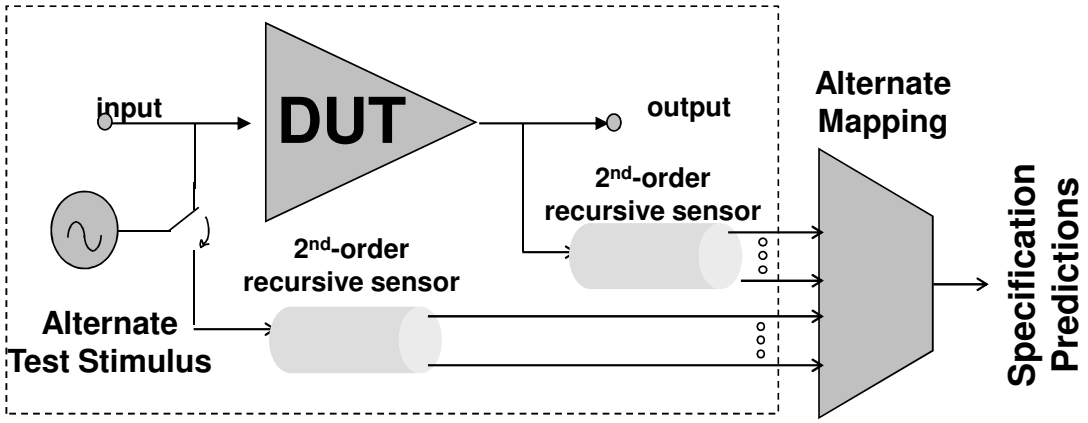


Figure 83: LNA experiment setup with recursive sensors.

The results of the experiments are depicted in Figure 84. First, an experiment with $3\sigma = nom/10$ is conducted with 25 instances in the training set, and the results are used as a reference for the experiments with recursive sensor and process monitor. The x-axis shows the ratio of the number of training samples used with the new sensor set to that of the reference set. The y-axis shows the ratio of rms error in 1dbComp prediction. The dashed line in the figure shows that without the process monitors, it is not even possible to generate a mapping with small training sets; even the training set that is 128 times larger than the reference is off-the-chart in terms of accuracy. On the other hand, the experiments with process monitors show good tracking at even small set sizes, where $4x$ represents a break point. Thus, when a 2^{nd} order process monitor is used with a 2^{nd} order recursive sensor, a training set that is 4 times larger is adequate to replace the 3125 times larger training set designed for process shifts. The ISO graphs for this configuration is given in Figures 85, 86, and 87.

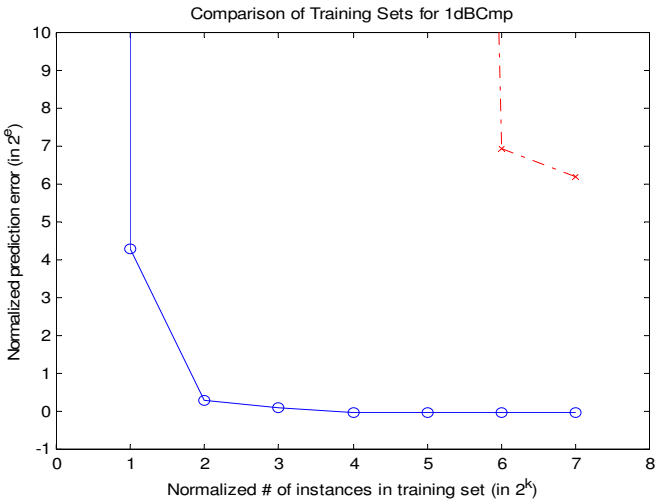


Figure 84: Prediction error versus size of training set. Dashed line shows the case with a 2^{nd}-order recursive sensor, solid line shows 2^{nd}-order process monitor and 2^{nd}-order recursive sensor.

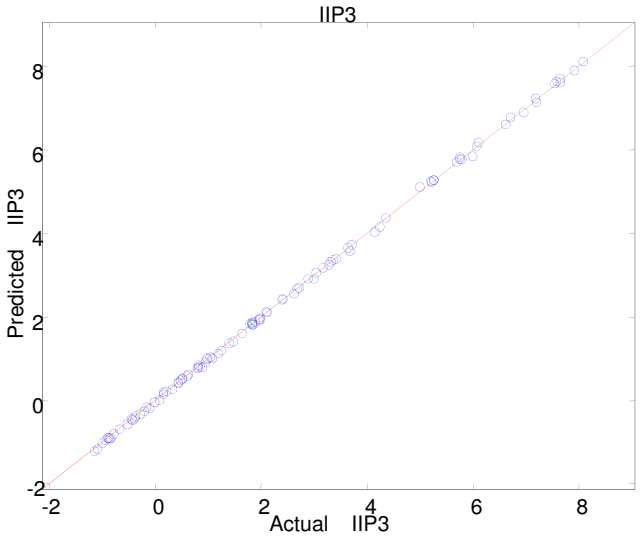


Figure 85: Predicted versus actual IIP3 with a 2^{nd} order process monitor and a 2^{nd}-order recursive sensor.

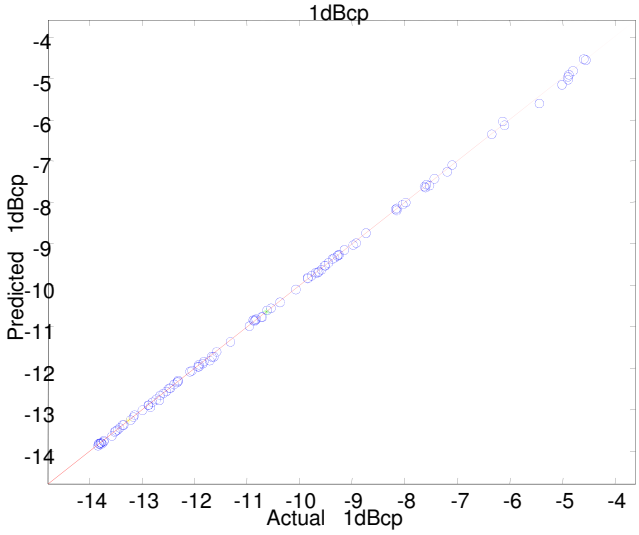


Figure 86: Predicted versus actual 1dBComp with a 2^{nd} order process monitor and a 2^{nd}-order recursive sensor.

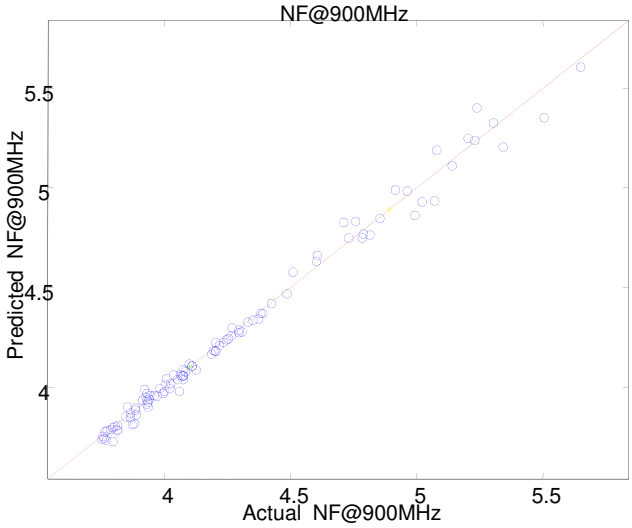


Figure 87: Predicted versus actual NF with a 2^{nd} order process monitor and a 2^{nd}-order recursive sensor.

Chapter V

DEFECT-BASED TESTING OF RF CIRCUITS

In this chapter, a methodology for efficient testing of RF devices with low-cost defect-based alternate tests is presented. Defect-based RF testing is a strong candidate for providing the best solution in terms of ATE complexity and cost. However, specification-based testing is still the norm for analog/RF because of the limitations of analog fault models. Unfortunately, as the amount of functionality packed into individual devices is increased with each generation, the cost of testing larger numbers of specifications also increases.

To address this challenge, the alternate testing methodology proposed in the previous sections, which significantly cuts costs associated with specification tests by crafting a single test stimulus and mapping the response signatures into all specifications at once, is modified for defect-based testing as well. In this chapter, we explore a new type of alternate testing that is more fundamental than defect-based or specification-based approaches. Rather than focusing on physical defect mechanisms or the way individual specifications are measured, *fault-based alternate testing* studies the abstractions of physical phenomena that cause specification violations. In this chapter, three different solutions are demonstrated featuring different trade-off points in testing cost, flexibility, and implementation complexity. All three methods guarantee that both defect-oriented and process-variation-oriented faults are detected equally well.

These methodologies can also make use of e-test data from wafer-probe to improve the quality of final product tests. This is indeed a defect-dictionary-independent test methodology with the reduced ATE requirements of defect-based test methods.

5.1 Fault Models for Defect-Based Alternate Testing

In the digital domain, effective test paradigms that rely on well-defined fault models, such as stuck-at faults, provide reliable, abstract and standard interfaces for powerful test algorithms to be developed, and for their performances to be compared relative to each other. Similar fault models for analog testing have been proposed; however, the degree of abstraction in these models has been mostly device specific, and consequently generic interfaces for test generation based on these models have not been developed. Although a generic fault-based test methodology is still the holy grail of analog testing, these devices are almost always tested in production for their performance specifications.

It is important to clarify some terms in the context of this study. First of all, a *specification* is a benchmark against which the performance of the DUT is measured; common RF specifications are gain, IIP3, NF, adjacent-channel power ratio (ACPR), and bit error rate (BER). Second, a *defect* is a physical phenomenon which creates a non-systemic and very local circuit anomaly, such as an open, a short or a bridge. Third, a *process variation* is the measurable difference of a physical manufacturing parameter caused by a physical deviation from the ideal manufacturing setup. This could be due to changes in the manufacturing environment, materials or methodology. The effects of such variations can be systemic within die, across wafer, or across lots of wafers. Such variations are caused by disturbances in the manufacturing process control mechanisms. Hence, even a tightly controlled process will have random distributions of manufacturing process parameters with almost constant expected values and standard deviations. Common process variations for CMOS can be observed in the oxide thickness and doping levels.

A *fault*, on the other hand, is an abstraction of physical phenomena which cause specification violations, such as a bridge between an output and an input pin, or a specific combination of process variations that reduce the bias current and causes

unacceptable gain. Note that, a *defect* or a moderate *process variation* will not always end up as a *fault*. They become a fault only when one or more specifications are violated. A defect is defined to be *redundant* if its impact on circuit performance does not manifest itself as a specification violation. This redundancy is especially valid for process variations because a good analog circuit designer tries to account for process distributions by corner and Monte-Carlo analysis. If large variations are realized only late in the production phase, where the design cannot be changed anymore, then the datasheet specifications are relaxed accordingly as long as the new minimal properties of the device have acceptable market value.

There are two more terms that need to be addressed. An *e-test* is a probe-level test applied to specific test structures on a wafer; it measures the electrical properties of these structures, such as the frequency of a ring oscillator. Last, a *process shift* is a change in the expected values of manufacturing process variables and occurs due to problems with process control. When e-test data obtained from measurements on affected wafers indicate that a process shift has occurred, feedback driven process compensation is performed to correct the problem for the next lot of wafers. Note that across-wafer process variations superposed on top of a process shift may result in some of the dice to be still fault-free. Again this presents a trade-off. If all dice are thrown away then this may have a large negative impact on yield; on the other hand, if all dice are packaged and tested, then valuable tester time is wasted mostly on faulty components.

5.2 A Survey of Test Paradigms

Existing fault models for analog/RF devices and associated coverage metrics are strongly dependent on the device and manufacturing technology. The common open, short and bridge models only cover a small subset of possible defects. These defect-oriented *catastrophic faults* seriously impair the functionality of the circuit. The larger

class, process-variation-oriented *parametric faults*, manifests itself as small deviations from the nominal circuit operating point. Most structural deviations in analog/RF circuits result in operating points well within acceptable limits. The differences between the many current analog test paradigms come from the way they address deviations; either performance deviations can be tested against specification limits or possible structural deviations can be eliminated against defect probabilities.

5.2.1 Defect-based testing

Defect-based testing, also called structural testing, [16–20, 22, 51, 133] works on the principle of negative elimination: each step of the test checks to see if a group of specific defects is present in the DUT; if the presence of a defect is detected, then the DUT is failed; otherwise it is shipped to the customer. The test generation scheme needs to have information about all kinds of possible defects, and will do better if provided with their probabilities. This information is manifested as a *defect-list* or a *defect-dictionary*. All catastrophic faults and large parametric faults are included in this dictionary. Process shifts are assumed to be detected by e-tests and all affected wafers are discarded; small parametric faults are assumed to be within acceptable performance limits by design or by proper redefinition of datasheet values.

The main advantage is that most defects can be easily detected with very simple and cheap equipment. The ideal ATE for a defect-based test system is "the voltmeter"; as long as "the voltmeter" reading stays within predefined safe levels while predetermined input patterns are applied, the device is fault-free. Although, applying a large set of these tests incurs long test time, multi-site testing can be applied to offset test time issues.

Defect-based testing as described above, requires extensive knowledge and characterization of the kinds of defects that can occur during silicon manufacturing. Also, if the defect-list is not accurate, or when new defects are introduced during production,

the test procedure needs to be redesigned. An additional problem with this approach is that simple voltmeter-type tests may not hold up if defects and large process variations *simultaneously* affect DUT performance. In such a scenario, simple tests may not detect performance excursions near the edge of performance acceptance caused by process variations.

The latter problem is significant because, first, in modern scaled CMOS processes, not all process variations can be effectively controlled anymore and relaxing performance acceptance limits to accommodate such large process variability can negate the benefits of technology scaling in the first place. Second, even in the presence of such large process variability, a significant percentage of dice still meet performance specifications. Throwing those dice out further sacrifices yield. These problems constitute the main reason why modern analog test methodologies focus on testing against specification limits rather than eliminating defect probabilities.

5.2.2 Specification-based testing

Specification testing of analog devices [21–25] is performed in a sequential manner mimicking individual characterization tests, but emphasizes throughput rather than completeness. Each step covers a single specification of the device: (i) the specific instruments on the ATE are selected and initialized, (ii) the selected instruments are routed to the appropriate DUT input and output ports by possibly using other coupling elements on the loadboard, (iii) the particular test stimulus is applied, (iv) the test response is digitized and the DUT performance value is calculated, and (v) the measured performance is compared against the desired specification. The process is repeated for every DUT performance specification.

As a result, the total testing time is linearly proportional to the number of datasheet specifications. As devices and their specifications become more complex

the increased testing cost dominates overall production cost. Similarly, extra functions often require additional measurement instruments to be integrated into the ATE, which increases the already prohibitive fixed cost [4], coming from signal integrity, precision and repeatability concerns.

5.2.3 Alternate Testing

In comparison to specification based testing, the alternate test methodology [42] is a different way of performing specification based testing. The alternate testing approach replaces the sequential nature of many different specification tests with a single test applied to the DUT, allowing the response signature to be mapped into all specifications at once. Hence, it cuts down testing time and reduces the instrumentation resource requirements of external ATE.

There are mainly two components to this flow: (i) *stimulus optimization* in which the test stimulus is carefully crafted to yield a significant correlation between the response and DUT performance variations, and (ii) *measurement synthesis* in which mapping functions from the DUT responses to their performance specifications are constructed using supervised learning [93]. The DUT performance parameters are not directly measured, instead the supervised learner predicts these values from the captured response signature. The test stimulus is engineered in such a way that the correlation between the DUT signatures and performance parameters is maximized given a target ATE cost and test time per DUT.

The stimulus optimization and measurement synthesis are performed on a sample set of training devices which constitute a representative set of process variations, shifts and defects present in the HVM environment. This way, the constructed functions can efficiently predict different specification values of a DUT given its signature response to the optimized alternate test stimulus [134]. The sample set can be selected from a number of hardware instances or can be generated by Monte-Carlo simulations.

5.2.4 Alternate Defect-Based Tests vs Classical Defect-Based Tests

Although case studies have shown that alternate test can provide significant improvement in test cost per component, accurate response capture instrumentation is still needed as the methodology eventually predicts the values of specifications using regression. The new methodology shown in this chapter allows robust test generation with more flexible constraints on ATE. It is inspired by the observation that traditional alternate tests [42] can detect very small deviations in the specifications of analog and RF devices. Hence, it is natural to conjecture that DUT responses to such tests have information to detect defects as well since these have a significant adverse effect of DUT performance. It also allows almost a continuous range of tradeoffs between the use of features from traditional specification-based testing methods and from defect-based testing.

The motivation of this work is to selectively replace some features of alternate test with those of defect-based test methods. By doing so, the ATE complexity and cost can be reduced by trading off alternate test complexity with parametric failure coverage with little or no impact on the coverage of physical defects. The resulting gamut of solutions have a common convergence point. The exact values of the specifications are not important although they may be obtained as a by-product of the alternate test procedure to varying degrees of accuracy. Redundant defects and minor process variations are not important. It is also not necessary to have a detailed fault-list for defect-based testing as long as the boundaries of specification violation under the *dominant* fault mechanisms are known. The methodology needs to figure out if a defect or process variation results in the violation of any one of the specifications. This is indeed the definition of a *fault*, hence we will classify these methods as *fault-based alternate testing techniques*. Three of these methods are studied in this work, (i) signature filter, (ii) alternate SVM, and (iii) dictionary-independent SVM. Each method represents a different trade-off point in testing cost, flexibility and implementation

complexity, and they all together provide a solution matrix for products with different market profiles. A total of 24 experiments are performed on a dataset of 1293 instances. Rest of the chapter describes the rationale behind this methodology, and explain several extensive experiments that study the trade-offs and limits of three methods. Also described is the constraint-driven alternate test stimulus generator (CodAlt), ATG formulated for defect-based alternate testing.

5.2.5 Failure Modeling with Defect-Based Alternate Test

The fault models for analog/RF devices and associated coverage metrics are device/product dependent. Furthermore, the common open, short and bridging models only cover a small portion of possible faults. These catastrophic faults seriously impair the functionality of the circuit. The larger class, called parametric faults, usually result from large process parameter variations and manifests itself as unacceptable deviations from the nominal circuit optimal operating point.

Figure 88 shows a Gaussian specification distribution for a two-sided analog metric. The discussion is easily adapted to single-sided analog specifications as well. The low and high specification limits shown in the figure define the acceptable range of specification values for the chosen device. This is under the assumption that the performance specification can be measured very accurately with no loss of precision. In general, this is not possible due to the uncertainties involved with low-cost test instrumentation. In the presence of measurement noise, the low and high test limit values are determined in such a way that any device with a value lower than the low test limit value or larger than the high test limit value is classified to be a "bad".

The ghost bands of Figure 88 are regions in which any device with an actual performance value within this range can be determined to be in the measurement tolerance bands with a statistically significant non-zero probability due to measurement noise. The measurement tolerance bands are bounded by the low and high test limits

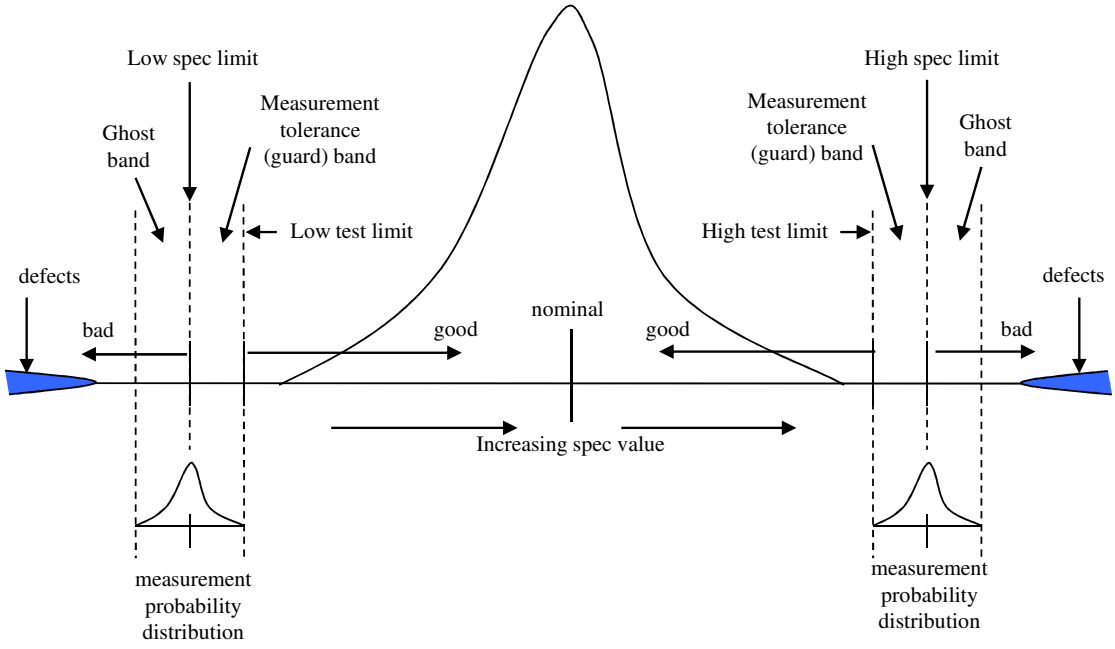


Figure 88: Parametric and catastrophic failure model.

on the sides closer to the nominal specification value. Conversely, any device with a specification value lower than the low limit has a statistically insignificant, practically zero, probability of having a measured performance value equal to or greater than the specification value corresponding to the low test limit in the presence of measurement noise. Similarly, any device with a performance value higher than the high spec limit has a practically zero probability of having a measured performance value equal to or less than the specification value corresponding to the high test limit. It is clear from the figure that the low and high test limits are set in such a way that some good devices may be classified as "bad" but no "bad" devices are classified as "good" due to measurement noise.

All manufactured devices map onto a point on the horizontal axis depending on their measured performance value. It is important to point out that in general, devices have multiple specifications and Figure 88 is a depiction of a multi-dimensional problem along a single-dimensional axis. The specific representation of the figure is

used for the sake of simplicity and can be extended easily to the general case. Figure 89 shows good, bad, and marginal devices along the specification axis.

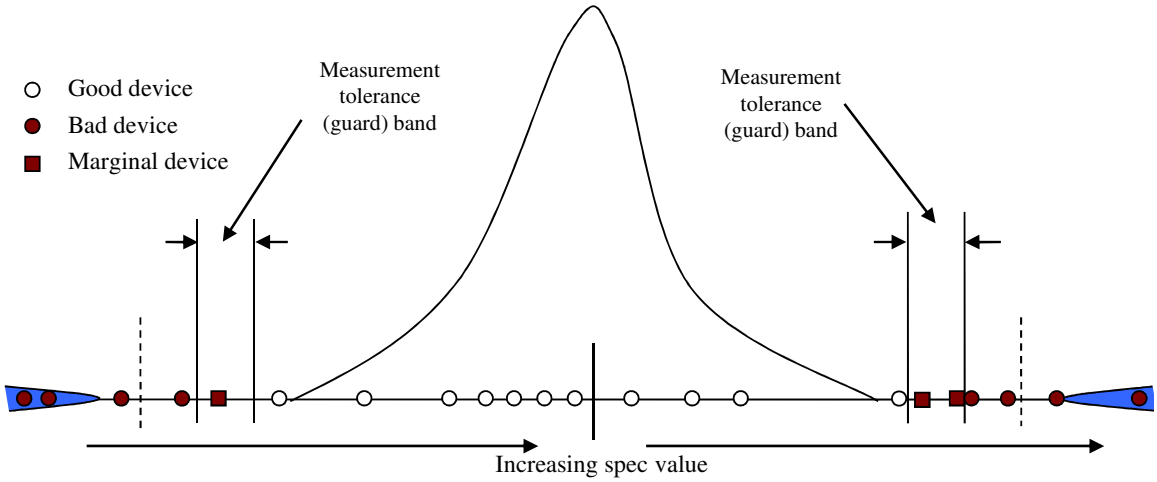


Figure 89: Good, bad, and marginal devices on a single-dimensional cross section of specification space.

A good device is defined as one that meets all its performance specifications. A bad device is defined to be one that does not meet at least one of its performance specifications. A marginal device is one that cannot be classified as being uniquely good or bad due to measurement noise. In industry, "good" devices as defined in 89, that are close to the test limit are also classified as "marginal" devices. Note that the definition of a bad or marginal device does not depend on whether it suffers from a parametric or catastrophic fault.

The following fault modeling approach is taken in this work. We assume that catastrophic faults will exhibit device behavior that is distinguishable from device behavior in the presence of parametric variations with a very high probability. In other words, the majority of the "marginal" devices, which are near or in the upper and lower measurement tolerance bands, are due to large parametric variations as opposed to physical defects. Consequently, we can define regions of the specification

axis for testing purposes as shown in Figure 90.

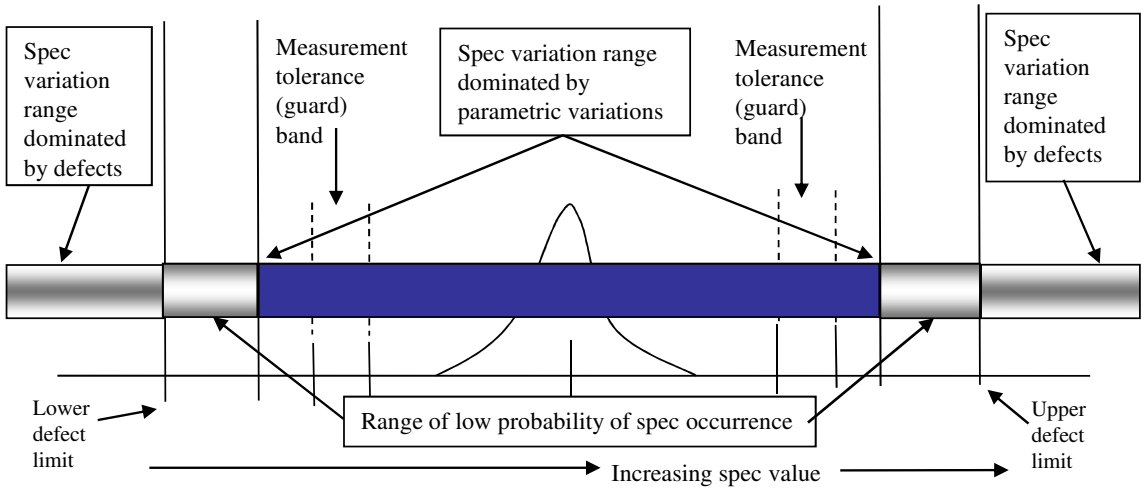


Figure 90: Regions of specification axis redefined for defect-based test purposes.

To formulate a viable strategy for defect-based testing, we first need to understand some core issues that have hindered development of defect-based test methods in the past. (a) Analog circuits can fail in almost infinite number of ways. Characterizing the many ways in which a particular circuit can fail in HVM is a difficult process and requires extensive collection of manufacturing test data across large numbers of wafers and extensive analysis of failure data. (b)Tests generated for a specified set of defects may not be valid for a defect mechanism that is discovered after the tests are generated. Hence, having a complete defect list is important prior to test stimulus generation. However, this is difficult to resolve for practical reasons.

To resolve problems (a) and (b), as the test of choice for defect discrimination, we propose using tests for the specification variation range dominated by parametric variations about the nominal as shown in Figure 90. The reasons for this are as follows: (i) Tests for the specification variation range dominated by parametric variations in the figure, do not depend on detailed defect-lists. Hence, tests, such as alternate

specification tests, can be generated based on knowledge of the defect-free circuit behavior under large parametric variations. There is only one way for a circuit to be defect-free, but thousands of ways for a circuit to be defective. These tests can predict circuit performance very accurately across the specification variation range dominated by parametric variations. (ii) Since the above alternate tests can track actual circuit specifications very accurately, it is reasonable to expect that these tests can easily detect defects that significantly perturb circuit behavior. Hence, intuitively, such tests should be easily able to detect defects that cause specification values to lie in the specification variation range dominated by defects. The few ICs that lie in the range of low probability of spec occurrence, can be classified as "marginal" or "bad" in this test approach. (iii) A key benefit of the above test approach is that the same alternate test stimulus can be used to determine if the circuit contains a defect, and also to determine the values of all of its test specifications from the obtained test response data if it is found to be defect-free. Hence, both the defect-based testing and parametric testing problems are solved with application of a single test stimulus.

The above approach, however, does not completely eliminate the problem of analyzing and simulating the impact of defects on DUT operation. As discussed later, it is necessary to know the lower and upper defect limits as shown in Figure 90 along the specification axis. If these limits move towards the nominal specification value, then distinguishing defects from parametric failures will become difficult. Even in this case, the proposed test approach will work with some minor modifications. On the other hand, if physical defects and parametric failures are distinguishable with a high probability, then it is necessary to know the values of these limits so that appropriate metrics can be developed for identifying defective ICs. Clearly, the further these limits are from the nominal, the easier it becomes to identify ICs with catastrophic defects.

The approach is streamlined in Figure 91. First, an alternate specification test

for the DUT is developed. This test is designed in such a way that the response of the DUT has strong statistical correlation with the test specification values of the DUT. Under this condition, the test specification values of the DUT can be predicted accurately from the obtained test response. The generated test does not require knowledge of defects or defect types, and hence can be generated relatively easily without extensive defect analysis. Information about tester digitizer resolution and measurement noise is used to design the test stimulus for maximum accuracy. Alternate test stimulus generation scheme described in previous chapters is used for this purpose. In the next step, the test is applied to the DUT and the response of the DUT is captured using a digitizer. Specific test response signatures are extracted from the test response for analysis. Such signatures consist of fourier or wavelet coefficients of the test response. These signatures are passed through a defect filter. The filter applies discrimination rules to the response signatures to determine whether a specific signature corresponds to a defect. Note that in our approach it is not necessary to know the signature corresponding to every possible defect type. In fact, we propose exactly the reverse. We try to determine if the obtained signature matches that of a defect-free circuit with very high probability. If this not the case, then the presence of a defect is indicated.

5.3 Experiment Setup

5.3.1 DUT and Figures of Merit

In order to demonstrate the proposed scheme, we use the single-ended 900MHz LNA, as shown here again in Figure 92. Seven specifications IIP3, 1dBCmp, NF, gain, RevIso, Sin, and Sout are simulated and recorded for every sample.

Each experiment is evaluated by three figures of merit based on the numbers of true positives (TP or T+), true negatives (TN or T-), false positives (FP or F+),

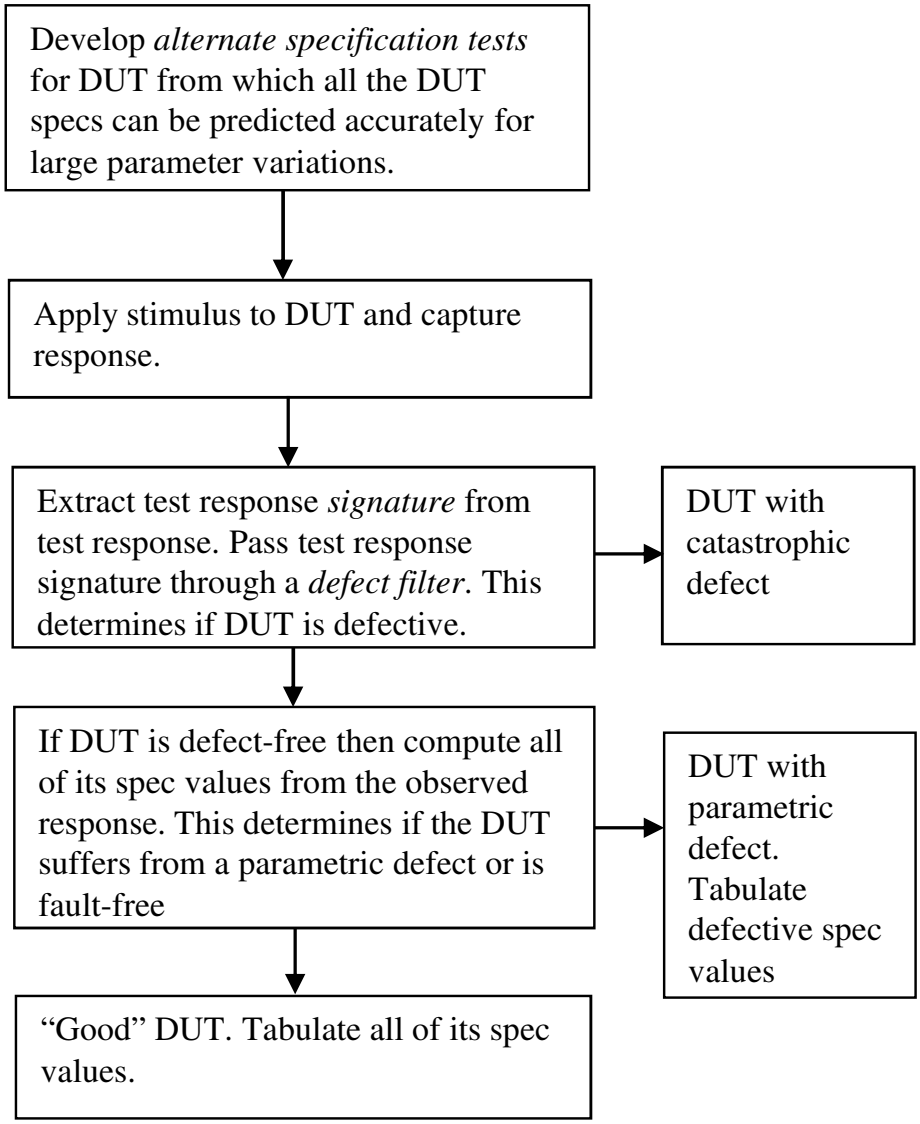


Figure 91: Defect-based alternate test approach.

and false negatives (FN or F-). A TP is a DUT that performs within the specification limits, and classified by the tester as a "passing" or "good" unit. A TN is a DUT, that performs outside the specification limits, and classified by the tester as a "failing" or "bad" unit. A FP is a DUT that performs outside the specification limits, but classified as a 'passing" or "good" unit. A FN is a DUT that performs within the specification limits, but classified as a "failing" or "bad" unit. An ideal test methodology only has TNs and TPs, with zero FN or FP. Practically, a certain

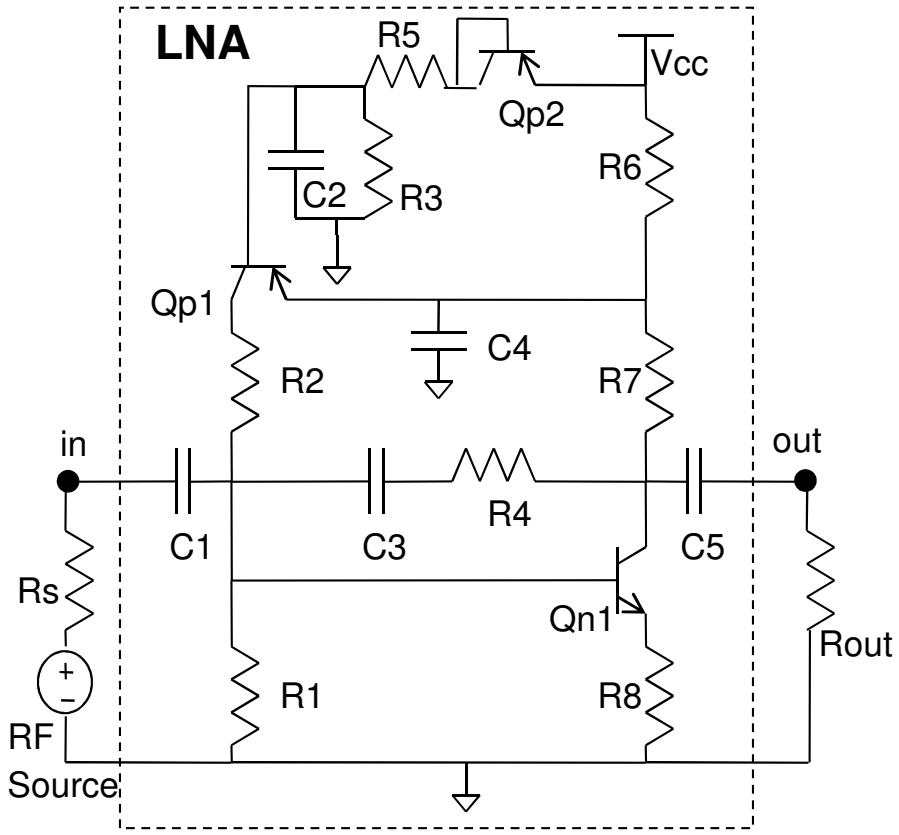


Figure 92: The schematic of the 900MHz LNA.

amount of FNs is acceptable. Higher the FNs, higher the negative yield impact will be. FPs are also called test escapes, and they are costly mistakes. The three figures of merit reflect preference for high yield and low test escapes, they are fault coverage (FC), pass coverage (PC), and accuracy (AC).

$$FC = \frac{TN}{TN + FP} \cdot 100\% \tag{23}$$

$$PC = \frac{TP}{TP + FN} \cdot 100\% \tag{24}$$

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\% \tag{25}$$

Complementary form of AC is the classification error (CE), which quantifies the percentage of misclassification:

$$CE = \frac{FP + FN}{TP + TN + FP + FN} \cdot 100\%. \tag{26}$$

A good test methodology yields 100% FC, no false positives, close to 100% PC, almost no false negatives, almost 100% AC, and almost 0% CE.

5.3.2 Test Stimulus Generation

Various alternate test stimulus generation algorithms are present in the literature [42, 63, 134], and as discussed in the previous chapters the proper implementation depends on both the structure of the device and the specifications under test. All three methods presented in this paper will follow the GA [87] based approach outlined in Section 3.3. First, each member of the 1293-instance set is simulated for target specifications. Once the actual specifications are calculated, the optimization loop starts. In each iteration, a generation in GA terms, there is a pool of candidate stimuli, coded into a gene notation, this is the current population of the GA. Since the DUT is an LNA, it is appropriate to limit the stimuli to frequency representations of multitone signals. A multitone can be encoded as a gene by associating the frequency value to a discrete gene location and assigning the digitized amplitude value to that storage as discussed in Section 3.4. The GA optimization starts with a random population of individuals. The goal is to evaluate a fitness value for each individual. Once these fitness values are calculated, the next generation of individuals are created. Three operations are used during this process, as described in Section 3.3: crossover combines parts of the genetic code of two parents to create an offspring, mutation randomly changes the value associated with a gene location, elite copies the genetic code to a new individual. Since individuals with higher fitness values have a higher chance to contribute to the next generation, each generation is more likely to have higher fitness values. Meanwhile, mutation introduces new traits and creates new combinations.

In this case study, the stimuli are limited to superpositions of two sinusoids only, their amplitudes and frequencies are constrained by the signal generator on the ATE.

Each generation is composed of 100 individuals and a typical GA process in this case takes around 80 generations to converge to an acceptable solution.

In every generation, each candidate stimulus is applied to every member of the set. The corresponding simulation responses are digitized with the given ADC accuracy, and converted into signatures. In this case, amplitudes of the dominant FFT components are the corresponding signatures. Only the signatures for the training set are used in creating the mapping functions. Then, the signatures for the validation set are fed into these mappings, and a fitness value is derived from the output. There are different ways to assign the fitness value, and the characteristics of the test stimulus generation methodology is mostly determined in these details. Next we will study each methodology in detail and see how the fitness values are defined. Note that, each methodology actually runs the optimization four times, each for a different set of e-test conditions.

5.4 Results

5.4.1 Signature Filter

The simplest extension of alternate testing to address faults is to divide the mapping process into two: first apply a coarse filter, called *defect filter*, to analyze the DUT response marking out catastrophic faults. Then, apply a mapping from the DUT response to the DUT specification values for parametric faults. The latter is easier for the stimulus generator as opposed to traditional alternate test because the test optimization space is smaller, smoother, and finer-grained. Note that this implementation does not need the e-test data to track process shifts. In this section, we show that a defect filter can be generated only from process variation data, without the need for a defect-dictionary.

When the signatures are compared against the defect filter, if any one of the signature components are outside the corresponding ranges defined by the filter, the

instance is labeled as a "fail", resulting in either an FN or an TN. Otherwise, numerical values for each specification are generated from the signatures. A prediction error (PE) is calculated for every specification of each instance. Note that an FP is to be avoided at all costs if possible and an FN acts as a yield penalty. For a TP instance, PE is the difference between predicted and actual specification values; for a TN instance, PE is zero; for an FN instance, PE is the largest PE among all TP or TN instances in the set; for an FP instance, PE is assigned a large penalty. All PEs are scaled by a nonlinear function described by:

$$EP_i = 1 + \frac{1 - \frac{PE_i}{PE_i^*}}{n} \quad for\ PE_i < PE_i^* \tag{27}$$

$$= \frac{PE_i^*}{PE_i} \quad for\ PE_i \geq PE_i^* \tag{28}$$

where n is the total number of specifications, and PE_i^* is the desired prediction error for the i^{th} specification, a realistic approximation to zero. The total scaled PE for an instance is given by $\sum_{i=1}^{n} EP_i$. Finally, the fitness value for a stimulus is the smallest EP among all instances.

The key step is the construction of the filter. Constructing a signature filter means finding upper and lower limits for each component of the signature under process parameter variations. There are basically two methods to determine these limits: (i) deterministically using corner analyses, or (ii) probabilistically using Monte-Carlo analyses. The deterministic way requires a sensitivity analysis tool which can generate precise corner data from the given netlist and process parameter variations. For high frequency RF components, the higher order correlations between process parameters require very complicated abstract models for efficient sensitivity analyzers. These abstract models can only be generated by a synthesis tool, which is most of the time not feasible for practical RF circuits. Hence, in this study, we employed a probabilistic method, which is also inherently compatible with alternate test generation process.

The probabilistic method uses the maximum and minimum values of the signature

components over the Monte-Carlo set generated for measurement synthesis. In order to account for the probabilistic corners, the limits are extended by a percentage of the corresponding range. The percentage value is selected to be 25% based on the characteristics of the supervised learner MARS. Furthermore, the upper limits of the responses corresponding to fundamental tones of the stimulus are extended by an additional margin, 3dB for the examples in this section. This way devices with exceptionally good operating points, which are designated by larger gain, are not penalized.

5.4.1.1 Example 1

As the signature filter method depends on the promise that no defect-dictionary is required, the first example studies a baseline sample set created with only process variations. No catastrophic faults or large R/C variations are introduced, and no defect filter is applied. This set of 400, S_{proc}, is generated by a simple Monte-Carlo simulation of 6 process variables: saturation currents (I_{SN},I_{SP}) and forward gains (V_{FBN},V_{FBP}) of NPN and PNP transistors, sheet resistance (R_{sheet}), and unit capacitance (C_{base}). The training set of 300, and the validation set of 100 are independently distributed, where process variables have linear $\pm$30% variation around their nominal values. To evaluate classification performance, artificial pass/fail boundaries are assigned in a way that 20% of the validation set fails the specification.

Figure 93 shows the IIP3 prediction and classification accuracy in an ISO graph. On the top-left corner, statistics about the specification are displayed, including nominal performance goal, pass/fail boundary, rms prediction error, maximum prediction error, and maximum prediction error percentage. On the bottom-right corner, statistics about classification is displayed, including FP and TN for this specification, and TN, TP, FN, and TP for all other specifications. The prefix "$\downarrow$" denotes the pictured specification, and prefix "$\exists$" denotes all other specifications. Yellow bars designate

pass/fail boundaries. There is no misclassification, which would otherwise be shown as red spots. Also, the prediction errors are smaller than typical measurement repeatability. Figure 94 shows the ISO graphs for the rest of the specifications. Overall, all instances are correctly classified with small prediction errors, as summarized in Table 5.4.1.1.

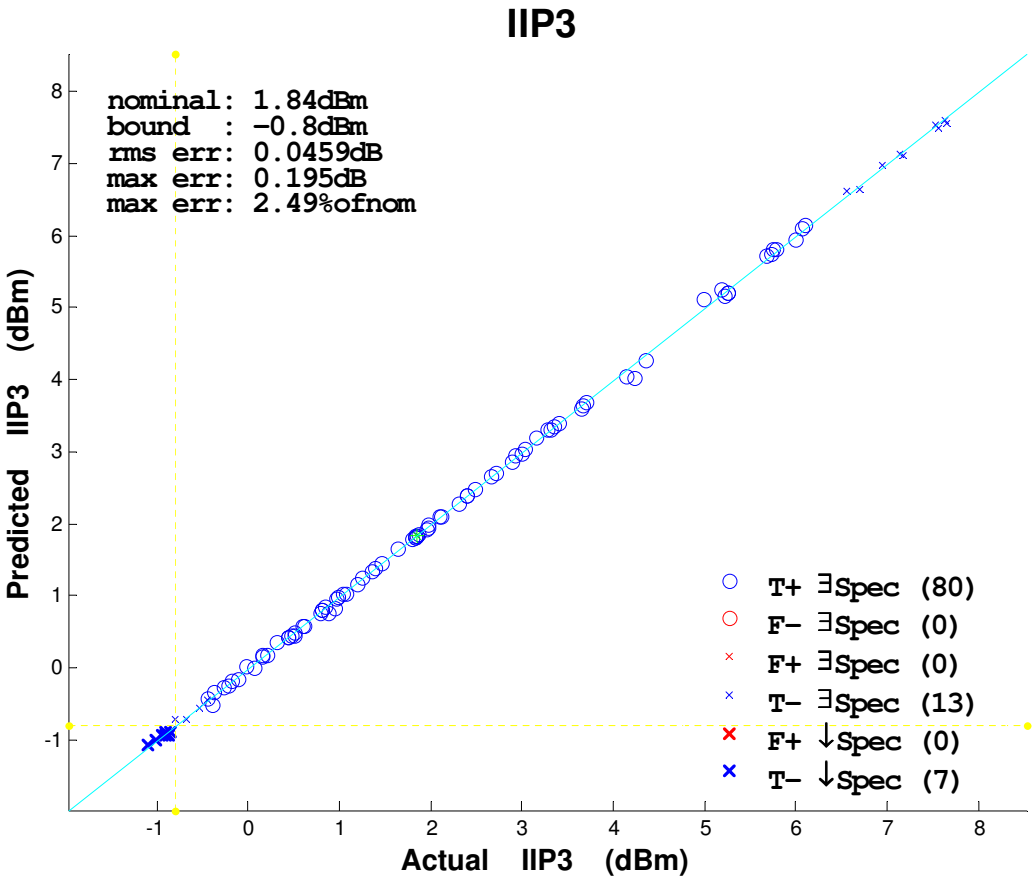


Figure 93: IIP3 prediction and classification accuracy on a sample set of parametric faults.

Table 12: Maximum and rms prediction errors for parametric faults.

Spec	**Unit**	**rms**	**max**
IIP3	dB	0.02	1.43
1dBC	dB	0.05	0.27
NF	dB	0.06	0.59
Gain	dB	0.06	0.42
RevIso	dB	0.02	0.09
Sin	/	0.01	0.04
Sout	/	0.01	0.06

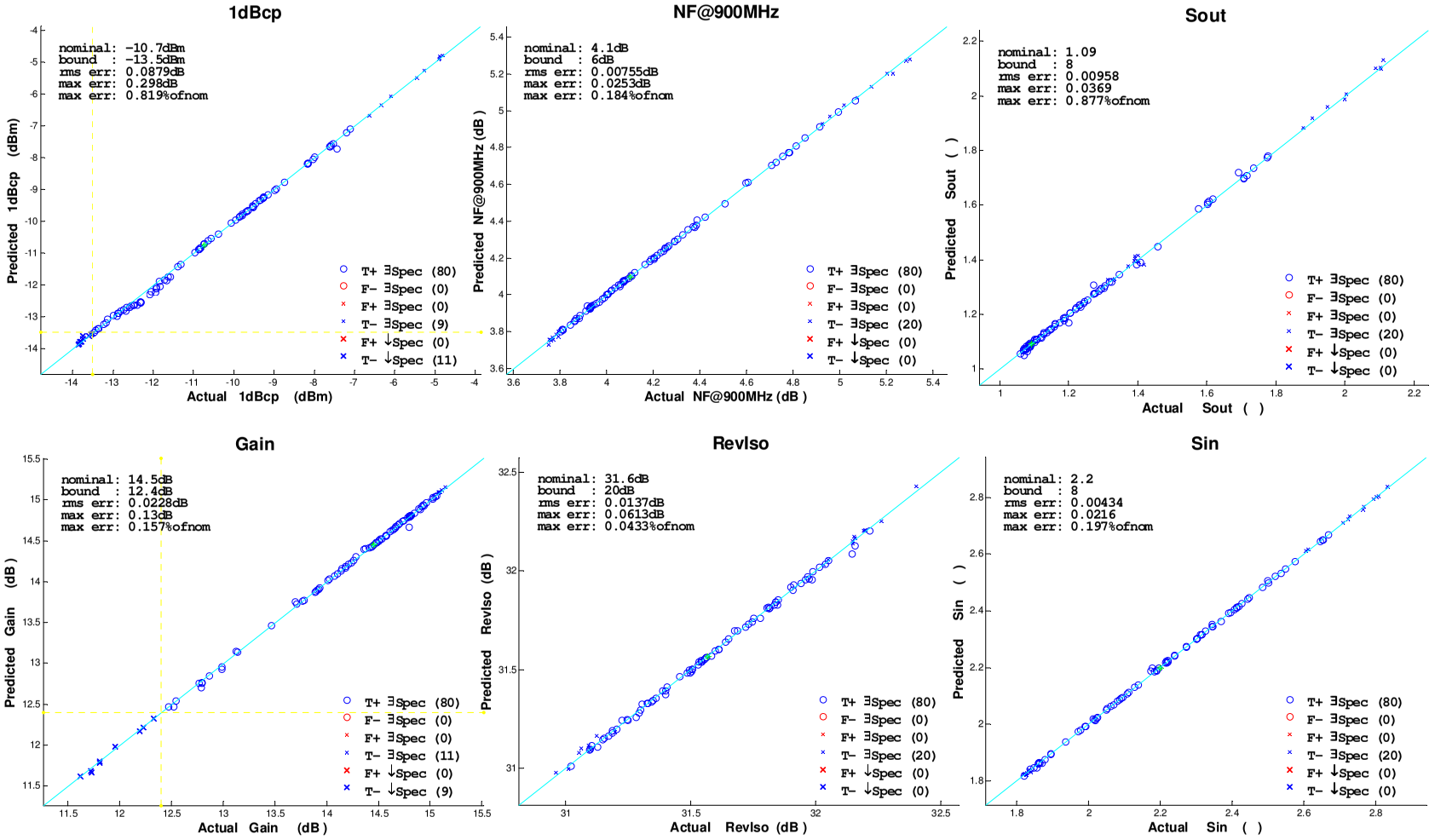


Figure 94: Prediction and classification accuracy on a sample set of parametric faults.

5.4.1.2 Example 2

In order to demonstrate the signature filter concept, the validation set in the first example is expanded with catastrophic faults, while keeping the training set exactly same. Using the same training set means the optimized alternate stimulus stays the same, and created with no access to a defect-dictionary. All possible opens, shorts and bridges are considered for catastrophic faults. Over 200 possible variations morph to 48 distinct faults. In addition, we have simulated five other fault modes, which represent ground bounces and resistive power paths. All faults are manually inserted into the netlist and devices are simulated for seven specifications under test, as shown in Figures 95 and 96. In order to cover a variety of catastrophic faults, we have simulated additional instances by replacing each resistor and capacitor with other ones at 1000-times, 10-times, 1.5-times, $1/1000^{th}$, $1/10^{th}$, and half of their original values, as shown in Figure 97. This way we introduced 78 extra fault modes, six for each capacitor and resistor value. All together, they add up to 131 candidates for catastrophic faults, *57 out of 131 are actually redundant faults*, meaning that all of their specifications are within the pass limits.

The signature filter is constructed as follows: (i) determine maximum and minimum values for each signature component over the 300-instance training set (dotted lines in Figure 98); (ii) enlarge the signature limits by $\pm 25\%$ of the range (solid lines in Figure 98); (iii) add an additional 3dB margin to the maximum limits for 900MHz and 920MHz fundamental responses. Table 5.4.1.2 lists the maximum/minimum values and filter upper/lower bounds for each signature component.

Three out of 57 redundant faults are eliminated directly by the signature filter.

```
27 C25 short in=net20
28 C25 open no input
29 R21 short no input & net20=ground
30 R21 open none
31 R22 short net9=ground
32 R22 open R22=100MOhm
33 C23 short net20=net24
34 C23 open no net24, C23, R6
35 R6 short net24=out
36 C21 open port0out=ground
37 C21 short out=port0out
38 R3 open no Q3, R3, net11
39 R3 short net11=net20
40 C24 open none
41 C24 short net39=ground
42 R0 open none
43 R0 short net29=out
44 C22 open none
45 C22 short no C22,R10 net27=ground
46 R10 open none
47 R19 open net14=ground
48 R19 short net14=net27
49 R23 open none
50 R23 short net39=power
51 Q4 shortBC net20=out no C23,R6
52 Q4 shortBE net20=net9
53 Q4 shortCE no Q4 out=net9
54 Q4 open no net9, R22, Q4
```

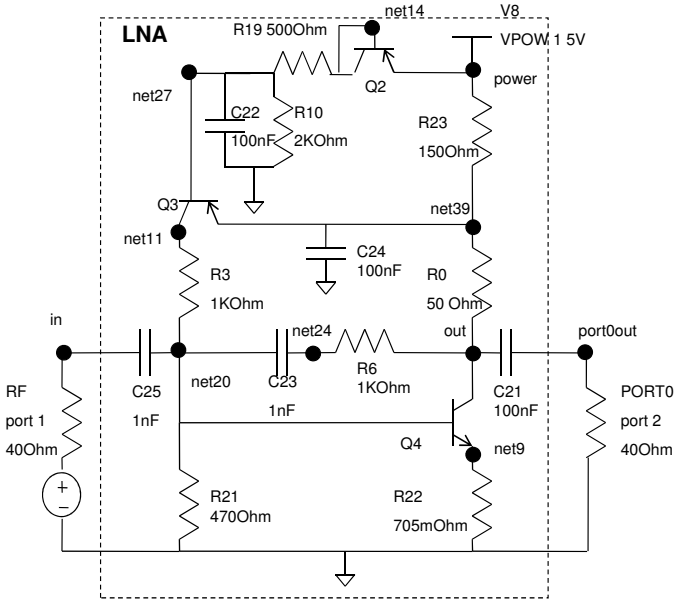


Figure 95: Opens, shorts, and bridge faults for the 900MHz LNA.

```
55 Q3 shortBC net11=net27
56 Q3 shortBE net39=net27
57 Q3 shortCE net39=net11 no Q3
58 Q2 shortCE net14=power no Q2
59 net bridge in=port0out
60 net bridge in-net20-net24-out-port0out
61 net bridge net20=net39
62 net bridge net11=out
63 net bridge net20=net27
64 net bridge net14=ground
65 net bridge net14=power no Q2
66 net bridge net27=power
67 net bridge net24=ground
68 net bridge net24=power
69 net bridge net11=power
70 net bridge net11=ground
71 net bridge out=power
72 net bridge out=ground
73 net bridge net20=ground no R21
74 net bridge power=ground no V8
75 vpow power 3V
76 vpow power 12V
77 ground power 3V
78 ground rail 10MOhm
79 ground rail 10kOhm
  (for ground bounces:
           Qs bulk stays ground
           sources stay ground )
```

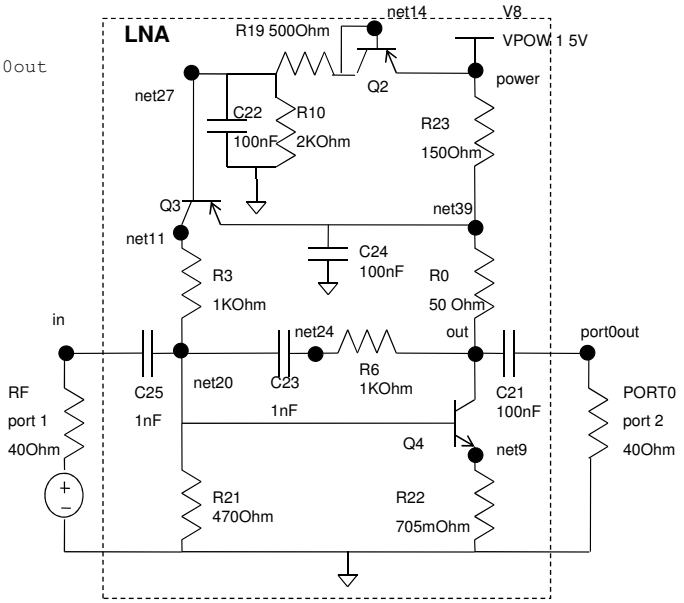


Figure 96: Opens, shorts, and bridge faults continued; ground bounce and resistive bridges.

	/1000	x1000	/10	x10	/2	x1.5
C25	1	2	80	81	106	107
R21	3	4	82	83	108	109
R22	5	6	84	85	110	111
C23	7	8	86	87	112	113
R6	9	10	88	89	114	115
C21	11	12	90	91	116	117
R3	13	14	92	93	118	119
C24	15	16	94	95	120	121
R0	17	18	96	97	122	123
C22	19	20	98	99	124	125
R10	21	22	100	101	126	127
R19	23	24	102	103	128	129
R23	25	26	104	105	130	131

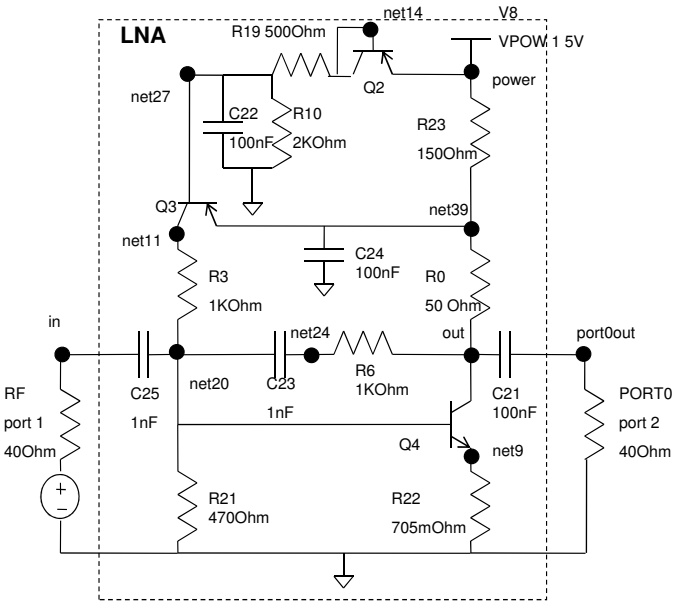


Figure 97: Potentially catastrophic R and C variations for the 900MHz LNA.

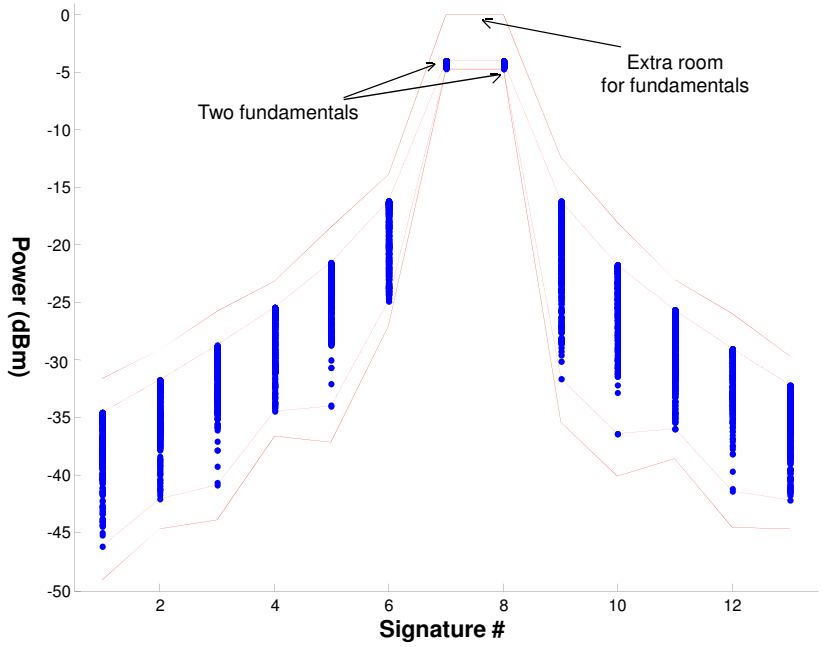


Figure 98: Signature filter; dotted lines show the maximum and minimum signature components while solid lines show signature filter upper and lower bounds after added margins.

Table 13: Signature filter upper and lower bounds

Frequency MHz	**max** dBm	**min** dBm	**lower** dBm	**upper** dBm
780	−34.56	−46.18	−49.08	−31.66
800	−31.67	−42.07	−44.67	−29.06
820	−28.73	−40.82	−43.85	−25.71
840	−25.40	−34.39	−36.64	−23.16
860	−21.55	−34.00	−37.12	−18.43
880	−16.15	−24.86	−27.04	−13.97
900	−4.032	−4.65	−4.81	−0.207
920	−4.060	−4.67	−4.82	−0.217
940	−16.23	−31.57	−35.40	−12.39
960	−21.69	−36.39	−40.06	−18.01
980	−25.60	−35.99	−38.58	−23.00
1000	−29.05	−41.43	−44.52	−25.95
1020	−32.19	−42.18	−44.68	−29.69

The rest of the redundant faults are passed into the alternate mappings for further classification, and only four are false negatives with no false positives. Overall, the alternate mappings with the signature filter scores 100% fault coverage on the 131-element catastrophic fault candidate set. Note that these results are obtained by training on the parametric fault set only, and without any fault list assumptions for catastrophic faults. Table 5.4.1.2 lists the rms and maximum prediction errors for the parametric fault set, the redundant fault set, and combined together. Figures 99a, 99b, 99c, 99d, 99e, 99f, and 99g show various types of redundant faults. The misclassifications are marked on the right side of the figures.

Figure 100 shows the ISO graph for IIP3 specification. The redundant faults are marked with rectangles. Figure 101 shows the ISO graphs for the rest of the

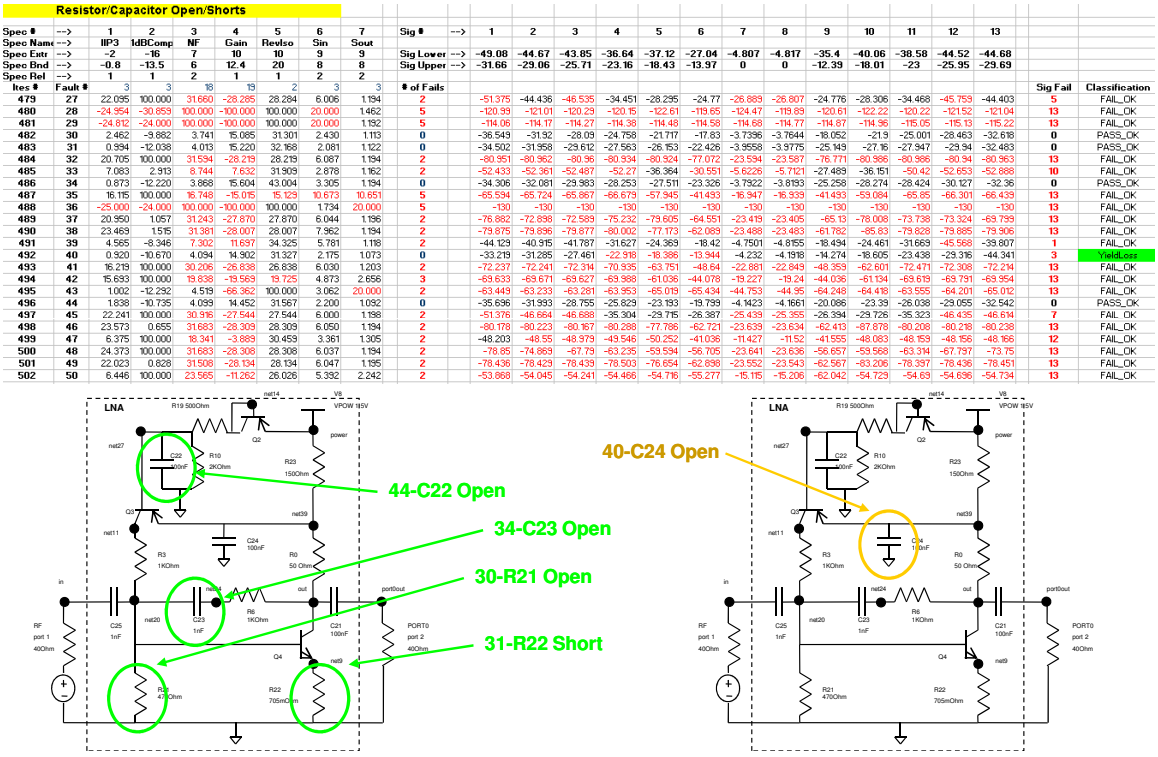


Figure 99: Classification of redundant faults; shown on the top, catastrophic faults introduced by resistor and capacitor open/shorts; on the left, correctly classified redundant faults; on the right, misclassified redundant faults (if any).

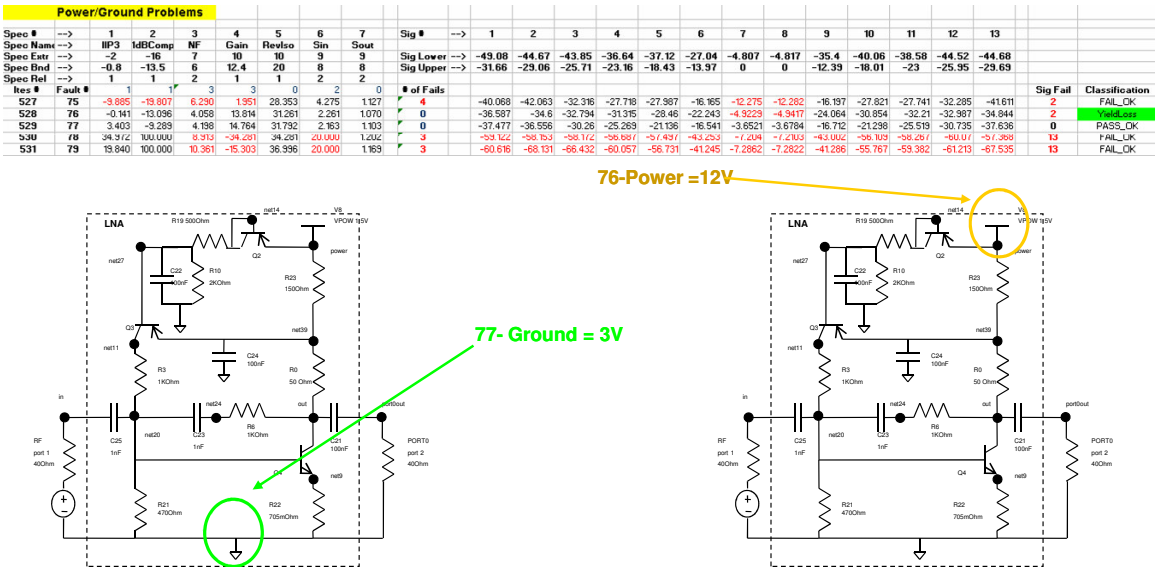


Figure 99: (cont'd) Catastrophic faults introduced by power or ground problems.

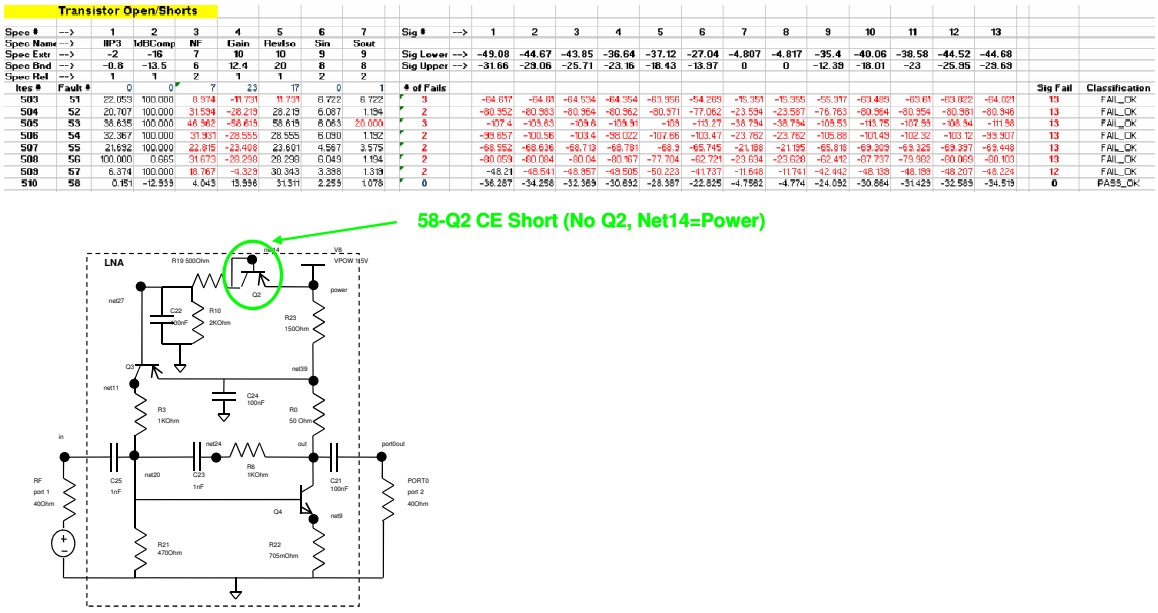


Figure 99: (cont'd) Catastrophic faults introduced by opens and shorts on transistors.

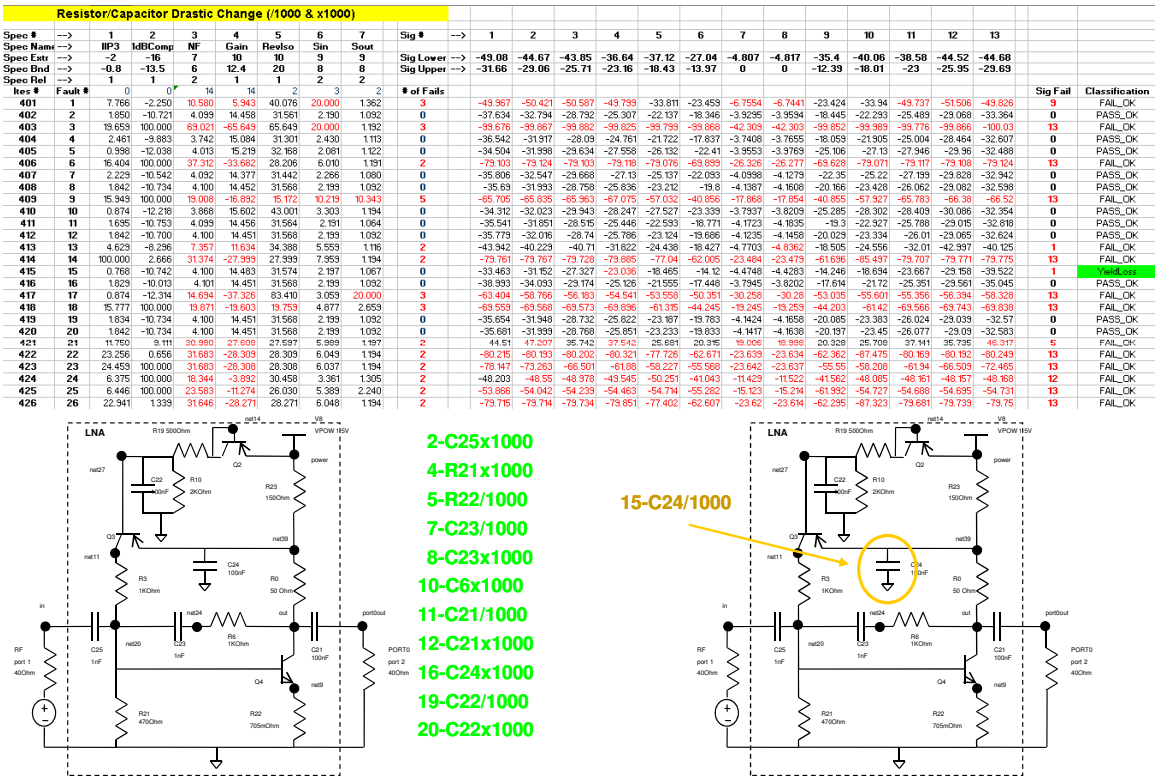


Figure 99: (cont'd) Catastrophic faults introduced by resistors and capacitors $x1000$ and $1/1000$ their original values.

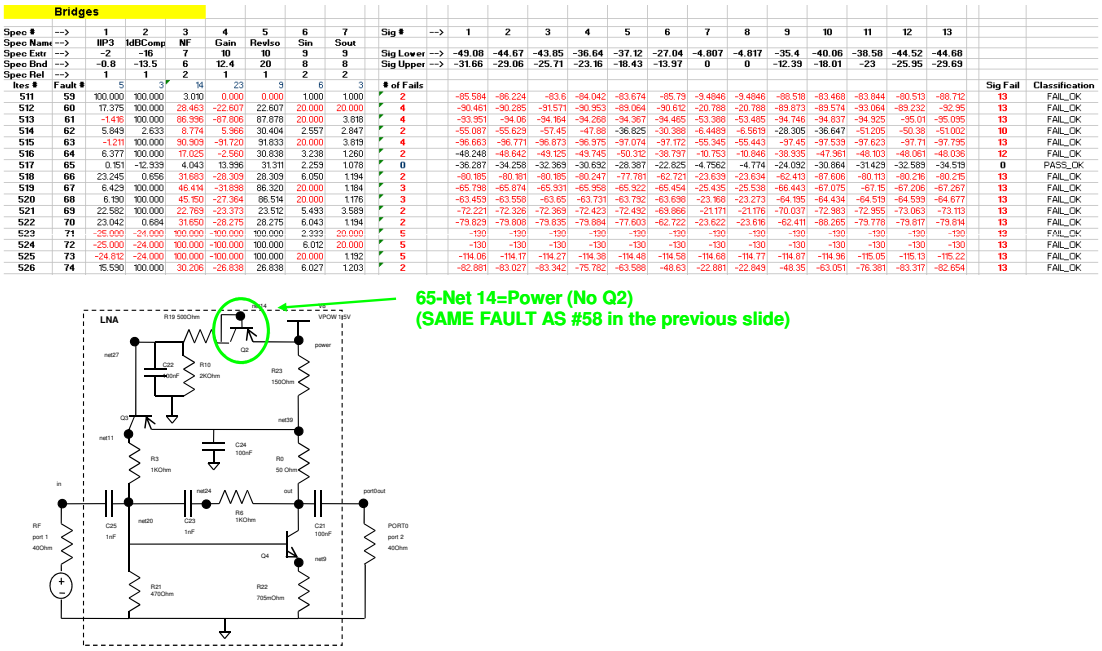


Figure 99: (cont'd) Catastrophic faults introduced by resistive bridges.

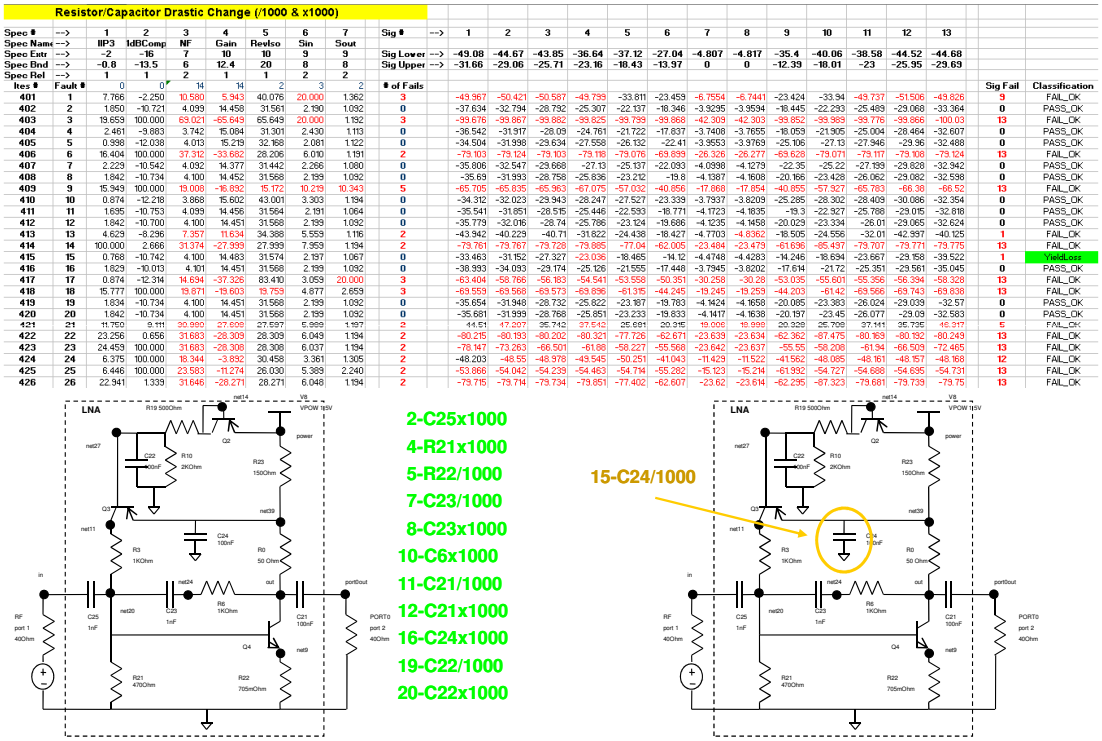
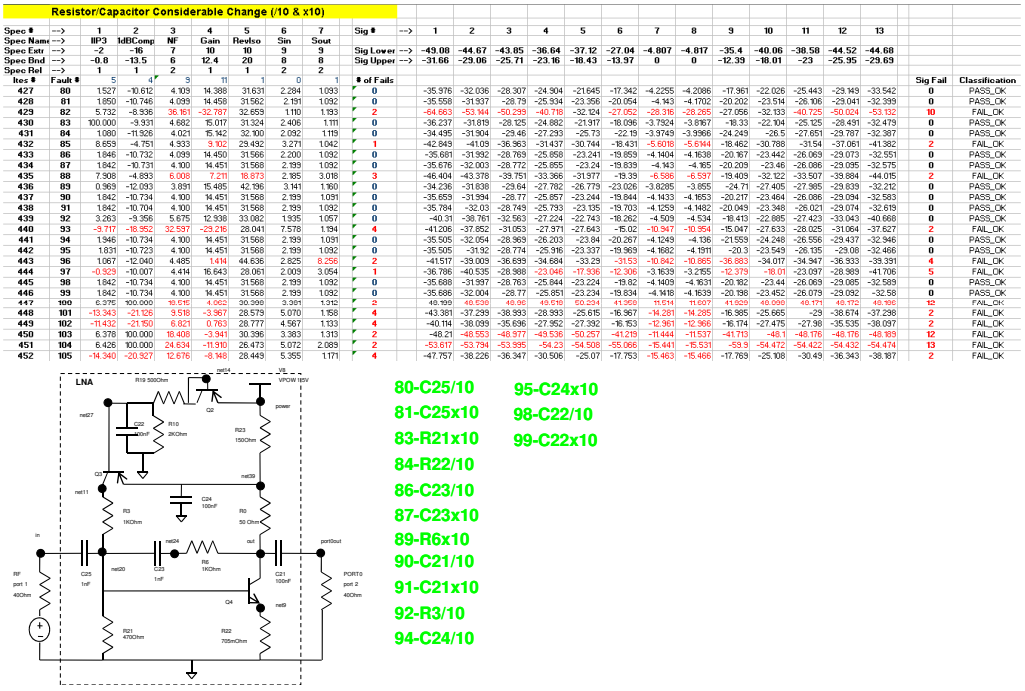


Figure 99: (cont'd) Catastrophic faults introduced by resistors and capacitors $x1000$ and $1/1000$ their original values.

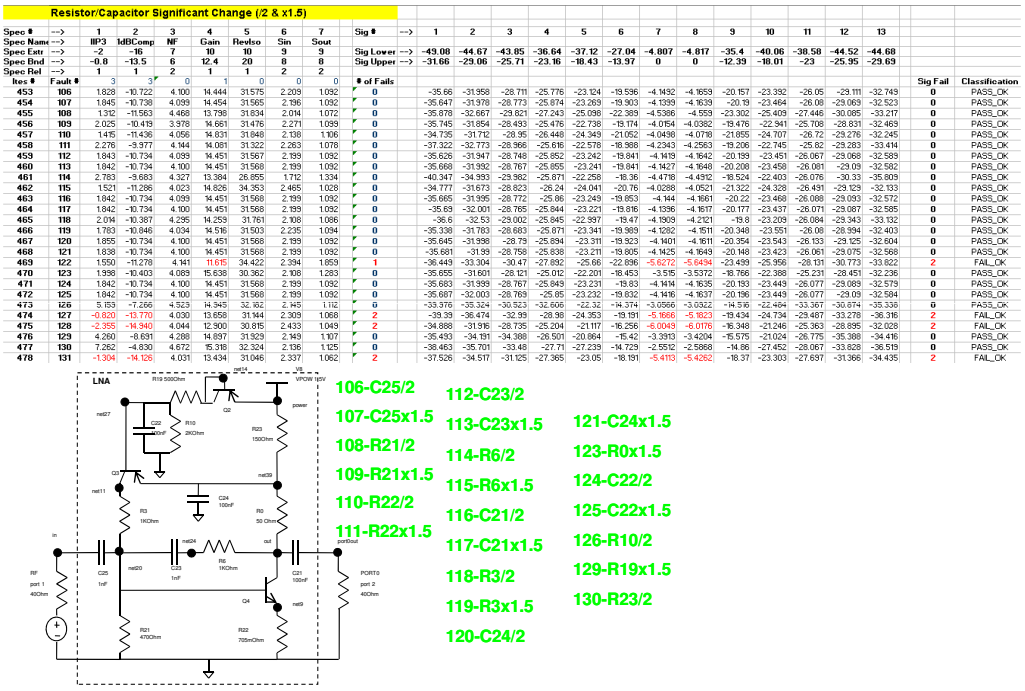


Resistor/Capacitor Considerable Change (/10 & x10)

Spec #	--->	1	2	3	4	5	6	7
Spec Name	--->	IIP3	IdBComp	NF	Gain	Revlso	Sin	Sout
Spec Extr	--->	-2	-16	7	10	10	9	9
Spec Bnd	--->	-0.8	-13.5	6	12.4	20	8	8
Spec Rel	--->	1	1	2	1	1	2	2
Ites #	Fault #	5	4	9	11	1	0	1

Sig #	--->	1	2	3	4	5	6	7	8	9	10	11	12	13		
Sig Lower	--->	-49.08	-44.67	-43.85	-36.64	-37.12	-27.04	-4.807	-4.817	-35.4	-40.06	-38.58	-44.52	-44.68		
Sig Upper	--->	-31.66	-29.06	-25.71	-23.16	-18.43	-13.97	0	0	-12.39	-18.01	-23	-25.95	-29.69		
# of Fails															Sig Fail	Classification
427	80	1.527	-10.612	4.109	14.388	31.631	2.284	1.093	0	-35.976	-32.036	-28.307	-24.904	-21.645	-17.342	-4.2255	-4.2086	-17.961	-22.026	-25.443	-29.149	-33.542	0	PASS_OK
428	81	1.650	-10.746	4.099	14.458	31.562	2.191	1.092	0	-35.558	-31.937	-28.79	-25.934	-23.356	-20.054	-4.143	-4.1702	-20.202	-23.514	-26.106	-29.041	-32.399	0	PASS_OK
429	82	5.732	-8.936	36.161	-32.787	32.653	1.110	1.193	2	-64.663	-53.144	-50.239	-40.718	-32.124	-27.052	-28.316	-26.265	-27.056	-32.133	-40.725	-50.024	-53.132	10	FAIL_OK
430	83	100.000	-9.331	4.682	15.017	31.324	2.406	1.111	0	-36.237	-31.813	-28.125	-24.882	-21.317	-16.096	-3.7924	-3.8767	-18.33	-22.104	-25.125	-28.491	-32.479	0	PASS_OK
431	84	1.080	-11.926	4.021	15.342	32.100	2.092	1.119	0	-34.495	-31.904	-29.46	-27.293	-25.73	-22.19	-3.9749	-3.9966	-24.249	-26.5	-27.651	-29.787	-32.387	0	PASS_OK
432	85	8.653	-4.751	4.933	9.102	29.432	3.271	1.042	1	-42.849	-41.03	-36.963	-31.437	-30.744	-18.431	-5.6018	-5.6144	-18.462	-30.788	-31.54	-37.061	-41.382	2	FAIL_OK
433	86	1.846	-10.732	4.099	14.450	31.566	2.200	1.092	0	-35.661	-31.932	-28.763	-25.858	-23.241	-19.859	-4.1404	-4.1638	-20.167	-23.442	-26.069	-29.073	-32.551	0	PASS_OK
434	87	1.842	-10.731	4.100	14.451	31.568	2.199	1.092	0	-35.676	-32.003	-28.772	-25.855	-23.24	-19.839	-4.143	-4.165	-20.209	-23.46	-26.086	-29.095	-32.575	0	PASS_OK
435	88	7.908	-4.893	6.008	7.211	18.873	2.185	3.018	3	-46.404	-43.378	-39.751	-33.366	-31.977	-19.39	-6.586	-6.597	-19.409	-32.122	-33.507	-39.884	-44.015	2	FAIL_OK
436	89	0.969	-12.093	3.891	15.485	42.196	3.141	1.160	0	-34.236	-31.838	-29.64	-27.782	-26.779	-23.026	-3.8285	-3.855	-24.71	-27.405	-27.985	-29.839	-32.212	0	PASS_OK
437	90	1.842	-10.734	4.100	14.451	31.568	2.199	1.091	0	-35.659	-31.994	-28.77	-25.857	-23.244	-19.844	-4.1433	-4.1653	-20.217	-23.464	-26.086	-29.094	-32.583	0	PASS_OK
438	91	1.842	-10.704	4.100	14.451	31.568	2.199	1.092	0	-35.784	-32.03	-28.749	-25.793	-23.135	-19.703	-4.1259	-4.1482	-20.049	-23.348	-26.021	-29.074	-32.619	0	PASS_OK
439	92	3.263	-9.356	5.675	12.938	33.082	1.935	1.057	0	-40.31	-38.761	-32.563	-27.224	-22.743	-18.262	-4.509	-4.534	-18.413	-22.885	-27.423	-33.043	-40.668	0	PASS_OK
440	93	-9.717	-18.952	32.537	-29.216	28.041	7.578	1.194	4	-41.206	-37.852	-31.053	-27.371	-27.643	-15.02	-10.947	-10.954	-15.047	-27.633	-28.025	-31.064	-37.627	2	FAIL_OK
441	94	1.946	-10.734	4.100	14.451	31.568	2.199	1.091	0	-35.505	-32.054	-28.963	-26.203	-23.84	-20.267	-4.1249	-4.136	-21.559	-24.248	-26.556	-29.437	-32.946	0	PASS_OK
442	95	1.831	-10.723	4.100	14.451	31.568	2.199	1.092	0	-35.505	-31.92	-28.774	-25.916	-23.337	-19.969	-4.1682	-4.1911	-20.3	-23.549	-26.135	-29.08	-32.466	0	PASS_OK
443	96	1.067	-12.040	4.485	1.414	44.636	2.825	8.256	2	-41.517	-39.005	-36.639	-34.884	-33.29	-31.53	-10.842	-10.865	-36.883	-34.017	-34.947	-36.933	-39.391	4	FAIL_OK
444	97	-0.923	-10.007	4.414	16.643	28.061	2.009	3.054	1	-36.786	-40.535	-28.388	-23.046	-17.936	-12.306	-3.1639	-3.2155	-12.379	-18.01	-23.097	-28.389	-41.706	5	FAIL_OK
445	98	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.688	-31.997	-28.763	-25.844	-23.224	-19.82	-4.1409	-4.1631	-20.192	-23.44	-26.069	-29.085	-32.589	0	PASS_OK
446	99	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.686	-32.004	-28.77	-25.851	-23.234	-19.834	-4.1418	-4.1639	-20.198	-23.452	-26.079	-29.092	-32.58	0	PASS_OK
447	100	6.375	100.000	10.516	4.062	30.309	3.361	1.312	2	46.199	46.536	46.06	46.016	50.234	41.959	11.514	11.607	41.929	46.098	46.171	46.172	46.196	12	FAIL_OK
448	101	-13.343	-21.126	9.518	-3.967	28.579	5.070	1.158	4	-43.381	-37.239	-38.933	-28.993	-25.615	-16.967	-14.281	-14.285	-16.985	-25.665	-29	-38.674	-37.236	2	FAIL_OK
449	102	-11.432	-21.150	6.821	0.763	28.777	4.567	1.133	4	-40.114	-38.039	-35.636	-27.952	-27.392	-16.153	-12.961	-12.966	-16.174	-27.475	-27.98	-35.535	-38.097	2	FAIL_OK
450	103	6.378	100.000	16.408	-3.941	30.396	3.383	1.313	2	-48.21	-48.553	-48.977	-49.536	-50.257	-41.219	-11.444	-11.537	-41.715	-48.1	-48.176	-48.176	-48.189	12	FAIL_OK
451	104	6.426	100.000	24.634	-11.910	26.473	5.072	2.089	2	-53.617	-53.794	-53.395	-54.23	-54.508	-55.066	-15.441	-15.531	-59.9	-54.472	-54.422	-54.432	-54.474	13	FAIL_OK
452	105	-14.340	-20.927	12.676	-8.948	28.449	5.355	1.171	4	-47.757	-38.226	-36.347	-30.506	-25.07	-17.753	-15.463	-15.466	-17.769	-25.108	-30.49	-36.343	-38.187	2	FAIL_OK

Figure 99: (cont'd) Catastrophic faults introduced by resistors and capacitors $x10$ and $1/10$ their original values.



Resistor/Capacitor Significant Change (/2 & x1.5)

Spec #	--->	1	2	3	4	5	6	7
Spec Name	--->	IIP3	IdBComp	NF	Gain	Revlso	Sin	Sout
Spec Extr	--->	-2	-16	7	10	10	9	9
Spec Bnd	--->	-0.8	-13.5	6	12.4	20	8	8
Spec Rel	--->	1	1	2	1	1	2	2
Ites #	Fault #	3	2	0	1	0	0	0

Sig #	--->	1	2	3	4	5	6	7	8	9	10	11	12	13		
Sig Lower	--->	-49.08	-44.67	-43.85	-36.64	-37.12	-27.04	-4.807	-4.817	-35.4	-40.06	-38.58	-44.52	-44.68		
Sig Upper	--->	-31.66	-29.06	-25.71	-23.16	-18.43	-13.97	0	0	-12.39	-18.01	-23	-25.95	-29.69		
# of Fails															Sig Fail	Classification
453	106	1.828	-10.722	4.100	14.444	31.575	2.209	1.092	0	-35.66	-31.958	-28.711	-25.776	-23.124	-19.596	-4.1432	-4.1653	-20.157	-23.392	-26.05	-29.111	-32.749	0	PASS_OK
454	107	1.845	-10.738	4.099	14.454	31.565	2.196	1.092	0	-35.647	-31.978	-28.773	-25.874	-23.263	-19.903	-4.1399	-4.1639	-20.19	-23.464	-26.08	-29.069	-32.523	0	PASS_OK
455	108	1.312	-11.563	4.468	13.798	31.834	2.014	1.072	0	-35.878	-32.667	-29.821	-27.243	-25.098	-22.389	-4.5386	-4.559	-23.302	-25.409	-27.446	-30.085	-33.217	0	PASS_OK
456	109	2.025	-10.419	3.978	14.661	31.476	2.271	1.099	0	-35.745	-31.854	-28.493	-25.476	-22.738	-19.174	-4.0154	-4.0382	-19.476	-22.941	-25.708	-28.831	-32.469	0	PASS_OK
457	110	1.415	-11.436	4.056	14.631	31.848	2.138	1.106	0	-34.735	-31.712	-28.95	-26.448	-24.349	-21.052	-4.0498	-4.0718	-21.855	-24.707	-26.72	-29.276	-32.245	0	PASS_OK
458	111	2.276	-9.977	4.144	14.081	31.322	2.263	1.078	0	-37.322	-32.773	-28.966	-25.616	-22.578	-18.988	-4.2343	-4.2563	-19.206	-22.745	-25.82	-28.283	-33.414	0	PASS_OK
459	112	1.843	-10.734	4.099	14.451	31.567	2.199	1.092	0	-35.626	-31.947	-28.748	-25.852	-23.242	-19.841	-4.1419	-4.1642	-20.199	-23.451	-26.067	-29.068	-32.589	0	PASS_OK
460	113	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.668	-31.992	-28.767	-25.855	-23.241	-19.841	-4.1427	-4.1648	-20.208	-23.458	-26.081	-29.09	-32.582	0	PASS_OK
461	114	2.763	-9.683	4.327	13.384	26.855	1.712	1.334	0	-40.347	-34.933	-29.982	-25.671	-22.258	-16.36	-4.4718	-4.4912	-18.524	-22.403	-26.076	-30.33	-35.809	0	PASS_OK
462	115	1.521	-11.286	4.023	14.826	34.353	2.465	1.028	0	-34.777	-31.673	-28.823	-26.24	-24.041	-20.76	-4.0288	-4.0521	-21.322	-24.328	-26.491	-29.125	-32.133	0	PASS_OK
463	116	1.842	-10.734	4.099	14.451	31.568	2.199	1.092	0	-35.665	-31.995	-28.772	-25.86	-23.249	-19.853	-4.144	-4.1661	-20.22	-23.468	-26.088	-29.093	-32.572	0	PASS_OK
464	117	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.69	-32.001	-28.765	-25.844	-23.221	-19.816	-4.1396	-4.1617	-20.177	-23.437	-26.071	-29.087	-32.585	0	PASS_OK
465	118	2.014	-10.387	4.295	14.259	31.761	2.108	1.086	0	-36.6	-32.53	-29.002	-25.845	-22.997	-19.47	-4.1909	-4.2121	-19.8	-23.209	-26.084	-29.343	-33.132	0	PASS_OK
466	119	1.763	-10.846	4.034	14.516	31.503	2.235	1.094	0	-35.338	-31.763	-28.683	-25.671	-23.341	-19.969	-4.1282	-4.1511	-20.348	-23.551	-26.08	-28.994	-32.403	0	PASS_OK
467	120	1.855	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.645	-31.998	-28.79	-25.894	-23.311	-19.923	-4.1401	-4.1611	-20.354	-23.543	-26.133	-29.125	-32.604	0	PASS_OK
468	121	1.838	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.681	-31.99	-28.758	-25.838	-23.211	-19.805	-4.1425	-4.1643	-20.148	-23.423	-26.061	-29.075	-32.568	0	PASS_OK
469	122	1.550	-11.278	4.141	11.615	34.422	2.394	1.859	1	-36.449	-33.304	-30.47	-27.892	-25.66	-22.696	-5.8272	-5.8494	-23.499	-25.956	-28.131	-30.773	-33.822	2	FAIL_OK
470	123	1.998	-10.403	4.089	15.638	30.362	2.108	1.283	0	-35.655	-31.601	-28.121	-25.012	-22.201	-18.453	-3.535	-3.5372	-18.766	-22.388	-25.231	-28.451	-32.236	0	PASS_OK
471	124	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.683	-31.999	-28.767	-25.849	-23.231	-19.83	-4.1414	-4.1635	-20.193	-23.449	-26.077	-29.089	-32.579	0	PASS_OK
472	125	1.842	-10.734	4.100	14.451	31.568	2.199	1.092	0	-35.687	-32.003	-28.769	-25.85	-23.232	-19.832	-4.1416	-4.1637	-20.196	-23.449	-26.077	-29.09	-32.584	0	PASS_OK
473	126	5.153	-7.266	4.523	14.345	32.192	2.145	1.112	0	-33.376	-35.324	-30.523	-32.606	-22.32	-14.374	-3.0566	-3.0322	-14.516	-22.464	-33.367	-30.674	-35.336	0	PASS_OK
474	127	-0.820	-13.770	4.030	13.658	31.144	2.309	1.068	2	-39.39	-36.474	-32.99	-28.98	-24.353	-19.191	-5.1666	-5.1823	-19.434	-24.734	-29.487	-33.278	-36.316	2	FAIL_OK
475	128	-2.355	-14.940	4.044	12.900	30.815	2.433	1.049	2	-34.888	-31.916	-28.735	-25.204	-21.117	-16.256	-6.0049	-6.0176	-16.348	-21.246	-25.363	-28.895	-32.028	2	FAIL_OK
476	129	4.260	-8.631	4.288	14.897	31.923	2.149	1.107	0	-35.493	-34.191	-34.388	-26.501	-20.864	-15.42	-3.3913	-3.4204	-15.575	-21.024	-26.775	-35.388	-34.416	0	PASS_OK
477	130	7.262	-4.830	4.672	15.318	32.324	2.136	1.125	0	-38.463	-35.701	-33.48	-27.71	-27.239	-14.729	-2.5512	-2.5868	-14.86	-27.452	-28.067	-33.828	-36.519	0	PASS_OK
478	131	-1.304	-14.126	4.031	13.434	31.046	2.337	1.062	2	-37.526	-34.517	-31.125	-27.365	-23.05	-18.191	-5.4113	-5.4262	-18.37	-23.303	-27.697	-31.366	-34.435	2	FAIL_OK

Figure 99: (cont'd) Catastrophic faults introduced by resistors and capacitors $1.5x$ and $1/2$ their original values.

Table 14: Maximum and rms prediction errors for parametric and catastrophic faults after signature filter.

		Parametric		Catastrophic		Both	
Spec	**Unit**	**rms**	**max**	**rms**	**max**	**rms**	**max**
IIP3	dB	0.02	1.43	0.57	1.54	0.36	1.54
1dBC	dB	0.05	0.27	0.55	1.39	0.33	1.39
NF	dB	0.06	0.59	0.18	0.47	0.12	0.59
Gain	dB	0.06	0.42	0.25	0.83	0.16	0.83
RevIso	dB	0.02	0.09	0.34	0.91	0.20	0.91
Sin	/	0.01	0.04	0.11	0.23	0.07	0.23
Sout	/	0.01	0.06	0.07	0.20	0.04	0.20

specifications. Overall, all instances are correctly classified with small prediction errors, as summarized in Table 5.4.1.2.

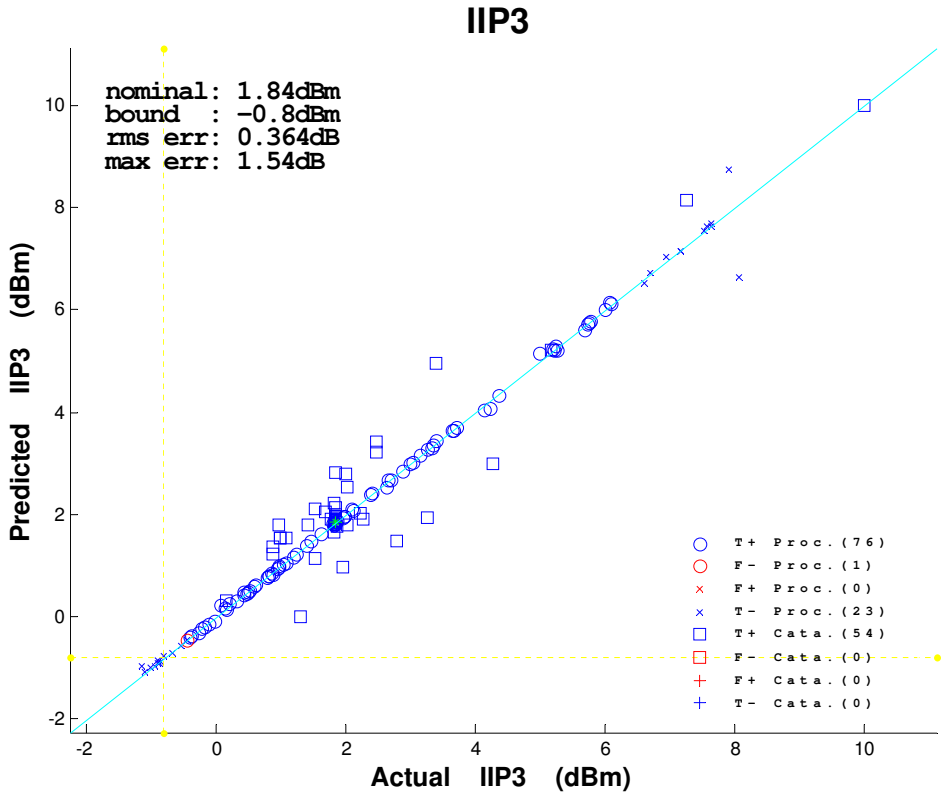


Figure 100: IIP3 prediction and classification accuracy with redundant faults.

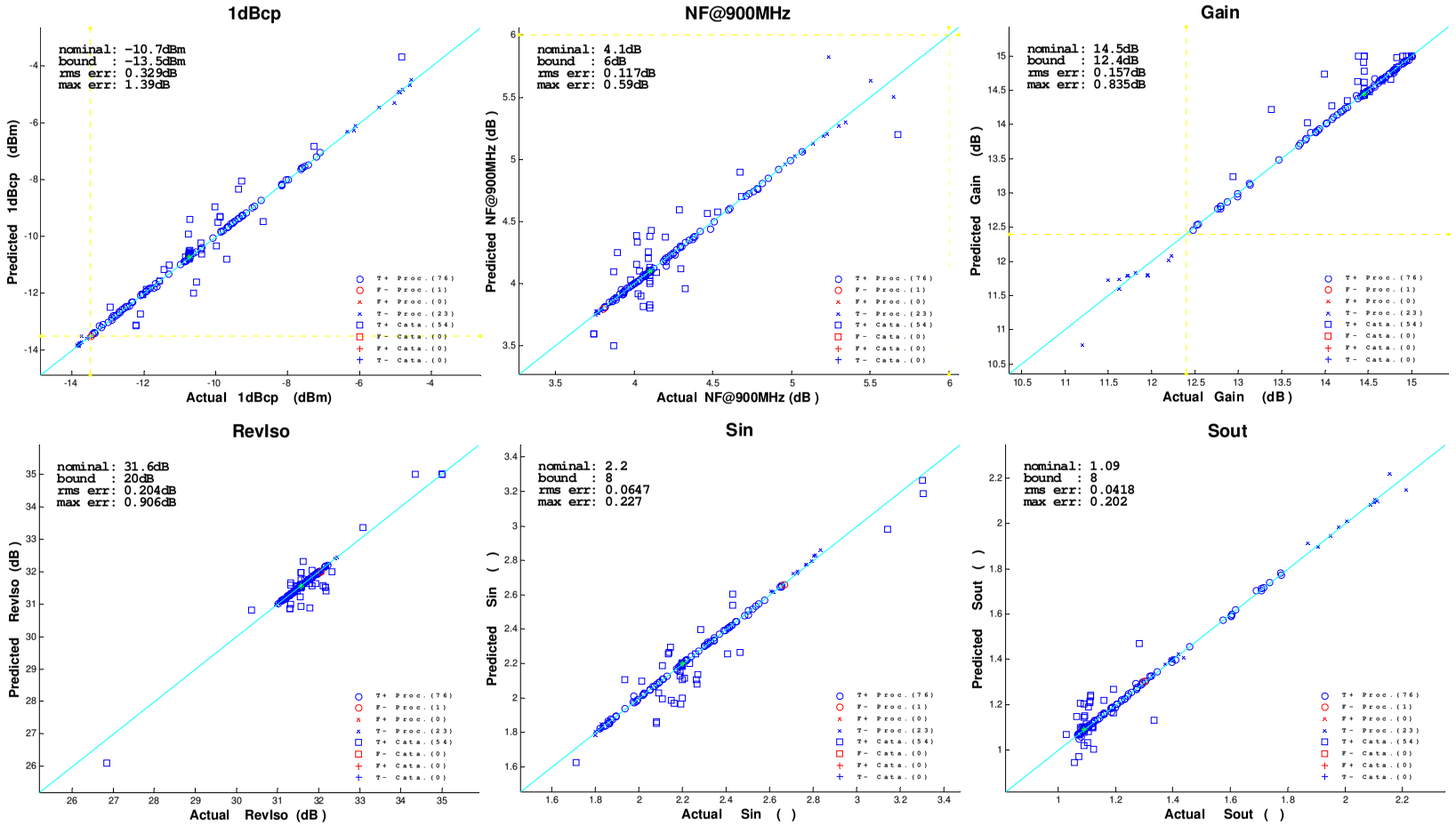


Figure 101: Prediction and classification accuracy with redundant faults.

5.4.1.3 Example 3

The next example extends the sample set to include instances resulting from a process shift. These new instances will help us evaluate the effectiveness of e-test data. A total of 1293 instances of the LNA are generated for this large dataset. In addition to the first set of 400, S_{proc}, a second set of 100, S_{sh1}, is generated by the same Monte-Carlo process in S_{proc} but the nominal operating point is shifted by significant differences on I_{SN} and V_{FBP}. This shift-set emulates the case where a process shift results in a nominal operating point far away from the original, all 100 instances fail either gain or NF specifications. A third set of 400 instances, S_{sh2}, is generated by the same Monte-Carlo process but this time with dominant process shifts on I_{SP} and V_{FBN}. This shift-set emulates a nominal operating point not far away from the original. In this case, only 166 instances fail for either IIP3 or 1dBCmp, the rest pass for all specs usually with better gain and NF. Figure 102 shows the histograms for gain and IIP3; green, blue, and magenta bars show the distributions of S_{proc}, S_{sh1}, and S_{sh2} respectively.

A fourth set of 393 instances, S_{def}, is generated by defects injected into the nominal operating points for S_{proc}, S_{sh1}, and S_{sh2}, 131 each as defined in the previous example. 162 of the set S_{def} are actually *redundant defects* meaning that all specifications are within the pass limits established by S_{proc}.

Given this dataset, 8 experiments are performed to explore three dimensions. The first dimension is the *ATE complexity*. The simulated responses are quantized with either 12 or 6 bits emulating effects of high or low-resolution ADCs. This is a realistic case study for RF testing, where a high-end mixed-signal ATE can be substituted for a more expensive RF ATE. This dimension can be used to explore the flexibility of different test methods by observing their performance differences in 12 and 6-bit configurations. *Availability of e-test data* is the second dimension. When disabled,

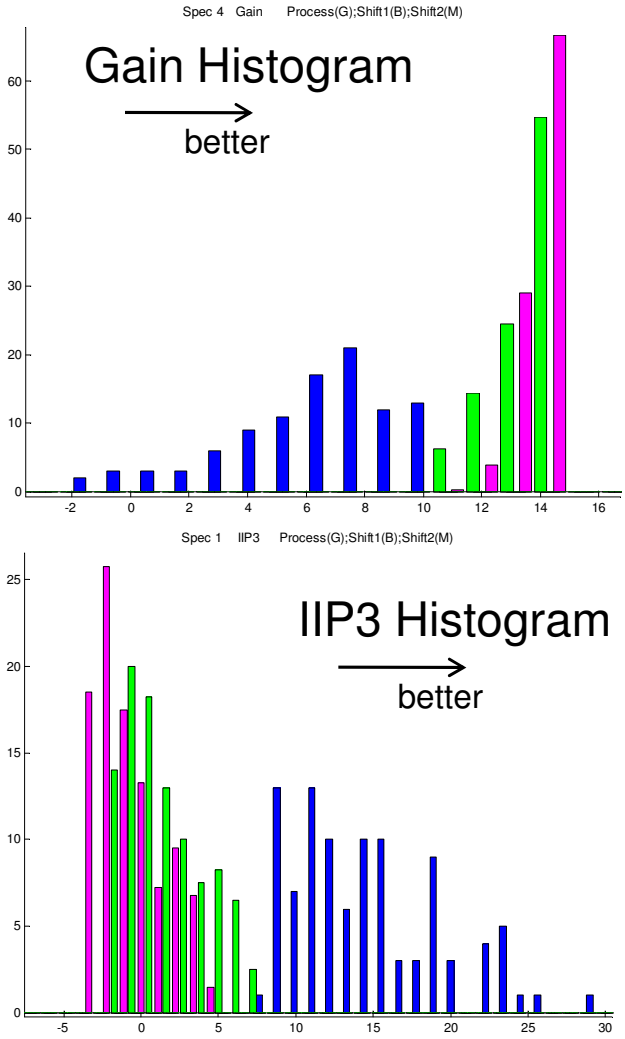


Figure 102: Histograms for gain (top) and IIP3 (bottom); green, blue, and magenta bars show the distributions of S_{proc}, S_{sh1}, and S_{sh2} respectively.

the alternate test methodology has one and only one optimized test stimulus for detecting faults. When enabled, they can make use of simple flags from wafer-probe testing and conditionally select from a set of previously optimized stimuli such that the selected stimulus best exploits the correlations in the subgroup designated by that e-test flag. In the given dataset, we assume that there are two exclusive flags to designate different shifts S_{sh1} and S_{sh2}, the absence of both flags means that the DUT belongs to S_{proc}; hence, each methodology can optimize and select from a total

of four different stimuli, one generic for e-test data disabled, and one specific for each of the three flag combinations when e-test data is enabled. For the sake of simplicity, the implementation of these specific e-tests is left out of the scope for this thesis. It is assumed that shifts can be exclusively flagged by wafer data. The final dimension is the *availability of a defect-dictionary*. Supervised learners learn by example from appropriate "training sets" of DUT populations with different performances. Their effectiveness, on the other hand, are evaluated on independent validation sets of DUTs selected by the cross-validation technique of [135]. When the defect-dictionary is enabled, any instances from S_{def} can be used as part of the training sets; otherwise, a series of experiments are performed each prohibiting a certain type of defect from appearing in the training sets, those defects only appear in validation sets, and the overall performance of the methodology is determined by averaging over the whole series.

The figures of merit for this method are listed in Table 5.4.1.3. There are two important observations here. First, the signature filter performs significantly better with the 12-bit ADC. The FC of 94% for the 6-bit ADC is unacceptable. This observation tells us that going after exact specification values requires significant precision. Consequently, this method may not be the best way to cut down ATE costs. Second, signature filter performs significantly better for 6-bit ADC when e-test data is available. FC moves up to 98% from 94%, and PC moves up to 93% from 88%. Although, 98% FC is still unacceptable, this observation brings a new angle to the previous claim: going after exact specification values causes the limited resources of the mapping function to spread thin over some regions of the measurement space and some FCs are let out from the tears; when e-test data is available and seperate mapping functions can be built for individual segments of the measurement space, it is as if the mapping function has three times the original resources. One minor observation is that the missing defect-dictionary makes no difference in figures of

merits (rows 3 and 4 compared with rows 1 and 2); this is expected because the limits of the signature filter are determined solely from Monte-Carlo simulations as opposed to instances with defects. With a twist, the availability of the defect-dictionary does not help improve the unacceptable FC for the 6-bit ADC experiments.

Table 15: Summary of Experiment 3 with the signature filter.

		Signature Filter					
		12-bit ADC			**6-bit ADC**		
		FC	**PC**	**AC**	**FC**	**PC**	**AC**
Non-restricted training set	w/o e-test	100	97	98	94	88	91
	w/ e-test	100	97	98	98	93	95
Restricted training set	w/o e-test	100	97	98	94	88	91
	w/ e-test	100	97	98	98	93	95

The experiment with restricted training set is expanded to be repeated with the sample set in Experiment 2. Figure 103 shows a detailed breakdown of evaluation metrics.

			Specification-Based Alternate Test + Signature Filter (MARS)													
			12-bit							6-bit						
			TN	TP	FN	FP	FC	PC	CE	TN	TP	FN	FP	FC	PC	CE
No Shift + Catast	Shift	0	0	0	0	0	100%	100%	0.000%	0	0	0	0	100%	100%	0.000%
	Process	400	0	400	0	0	100%	100%	0.000%	0	368	32	0	100%	92%	8.000%
	Catast	131	70	54	7	0	100%	89%	5.344%	67	46	15	3	96%	75%	13.740%
	All	531	70	454	7	0	100%	98%	1.318%	67	414	47	3	96%	90%	9.416%
Shift + Catast w/o Etest	Shift	500	266	231	3	0	100%	99%	0.600%	248	212	22	18	93%	91%	8.000%
	Process	400	0	400	0	0	100%	100%	0.000%	0	372	28	0	100%	93%	7.000%
	Catast	393	210	162	21	0	100%	89%	5.344%	201	138	45	9	96%	75%	13.740%
	All	1293	476	793	24	0	100%	97%	1.856%	449	722	95	27	94%	88%	9.435%
Shift + Catast w Etest	Shift	500	266	234	0	0	100%	100%	0.000%	256	215	19	10	96%	92%	5.800%
	Process	400	0	400	0	0	100%	100%	0.000%	0	381	19	0	100%	95%	4.750%
	Catast	393	210	162	21	0	100%	89%	5.344%	210	162	21	0	100%	89%	5.344%
	All	1293	476	796	21	0	100%	97%	1.624%	466	758	59	10	98%	93%	5.336%
			True Negative - Fail OK	True Positive - Pass OK	False Negative - Yield Loss	False Positive - DPM	Fault Coverage - % of Correct Negative Pred - TN/TN+FP	Redundant Coverage - % of Correct Redundant Pred - TP/TP+FN	Accuracy - % of Correct Pred - TN+TP/ALL	True Negative - Fail OK	True Positive - Pass OK	False Negative - Yield Loss	False Positive - DPM	Fault Coverage - % of Correct Negative Pred - TN/TN+FP	Redundant Coverage - % of Correct Redundant Pred - TP/TP+FN	Accuracy - % of Correct Pred - TN+TP/ALL

Process Parameter Variations + Catastrophic Faults + Large Parametric Faults
Process Par + Catast + Large Param 2 large process parameter shifts No e-test data assumed
Process Par + Catast + Large Param 2 large process parameter shifts E-test data used for stimulus and classifier selection

Figure 103: Signature Filter; detailed breakdown of experiments 2 and 3 with restricted training set in terms of evaluation metrics.

5.4.2 Alternate Tests with SVM

The second method is to abandon specification prediction and simply stick with a go/no-go decision. This can be achieved by using classifiers instead of regression methods. Although similar solutions have been proposed in [136,137], these methods do not provide a way for automated stimulus generation, but rather deal with crafting the best decision boundary given a test stimulus. In this section, we demonstrate a test generation algorithm based on *SVM* [135,138]. The two experiments performed show that (i) the go/no-go approach has a significant effect on ATE complexity when compared to the regression-based mapping approach (use of 6-bit vs 12-bit ADC), and (ii) e-test data can be used to improve the performance of this approach in terms of fault detectability. On the other hand, this method has a strong dependence on the defect-dictionary, the training set of samples needs to reflect a fair distribution of expected faults.

This method implements a classifier as the mapping function, hence the output is either a "pass" or a "fail", the specification values are not predicted at all. The supervised SVM learner is implemented with WEKA tools [139]. This time a single classification error (CE) replaces the PE in Section 5.4.1. Each FN adds one to the CE, while each FP adds a hundred. The CE is scaled by a nonlinear function:

$$EC = 1 + \frac{1 - \frac{CE}{CE^*}}{n} \quad for\ CE < CE^* \tag{29}$$

$$= \frac{CE^*}{CE} \quad for\ CE \geq CE^* \tag{30}$$

where n is the total number of specifications, and CE^* is the desired classification error, a realistic approximation to zero. EC serves as the fitness value for the stimulus.

Experiment 3 in the previous section is repeated with this classifier. The figures of merit for this method are listed in Table 5.4.2. When compared with the 6-bit/12-bit results for the signature filter method, it is obvious that alternate SVM presents a much more flexible methodology. The FC of 98% for the 6-bit ADC is close to the

acceptable range. Although PC is still low at 85%, this is mainly due to the 1-to-100 penalty of FNs versus FPs. If desired, a more balanced solution can be achieved by changing the penalty ratio in favor of FNs. The highlight of these experiments is that when e-test data is available, it is possible to achieve the figures of a 12-bit ADC with the 6-bit ADC (FC 100%, PC 99%, AC 99%). *This observation tells us that going after simple go/no-go decisions -rather than precise specification predictions- provides enough flexibility to cut down ATE costs.* However, one must not forget that this method heavily depends on the availability of the defect-dictionary. When certain defect classes are absent from the training sets the figures significantly go down (FC 97%), even in the precense of e-tests as seen in the last row.

Table 16: Summary of Experiment 3 with alternate SVM.

		Alternate SVM					
		12-bit ADC			**6-bit ADC**		
		FC	**PC**	**AC**	**FC**	**PC**	**AC**
Non-restricted	w/o Etest	100	93	95	98	85	90
training set	w/ Etest	100	99	99	100	99	99
Restricted	w/o Etest	97	86	87	94	79	88
training set	w/ Etest	97	93	92	97	93	93

The experiment with restricted training set is expanded to be repeated with the sample set in Experiment 2. Figure 104 shows a detailed breakdown of evaluation metrics.

			Fault-Based Alternate Test (SVM)																	
			12-bit									6-bit								
			TN	TP	FN	FP	FC	PC	CE	RE	PR	TN	TP	FN	FP	FC	PC	CE	RE	PR
No Shift + Catast	Shift	0	0	0	0	0	100%	100%	0.000%	-	-	0	0	0	0	100%	100%	0.000%	-	-
	Process	400	0	400	0	0	100%	100%	0.000%	100%	100%	0	400	0	0	100%	100%	0.000%	100%	100%
	Catast	131	70	59	2	0	100%	97%	1.527%	46%	100%	70	57	4	0	100%	93%	3.053%	45%	100%
	All	531	70	459	2	0	100%	100%	0.377%	87%	100%	70	457	4	0	100%	99%	0.753%	87%	100%
Shift + Catast w/o Etest	Shift	500	266	203	31	0	100%	87%	6.200%	43%	100%	266	187	47	0	100%	80%	9.400%	41%	100%
	Process	400	0	378	22	0	100%	95%	5.500%	100%	100%	0	372	28	0	100%	93%	7.000%	100%	100%
	Catast	393	210	177	6	0	100%	97%	1.527%	46%	100%	201	138	45	9	96%	75%	13.740%	41%	94%
	All	1293	476	758	59	0	100%	93%	4.563%	61%	100%	467	697	120	9	98%	85%	9.977%	60%	99%
Shift + Catast w Etest	Shift	500	266	234	0	0	100%	100%	0.000%	47%	100%	266	233	1	0	100%	100%	0.200%	47%	100%
	Process	400	0	399	1	0	100%	100%	0.250%	100%	100%	0	399	1	0	100%	100%	0.250%	100%	100%
	Catast	393	210	177	6	0	100%	97%	1.527%	46%	100%	201	175	8	0	100%	96%	2.083%	47%	100%
	All	1293	476	810	7	0	100%	99%	0.541%	0.63	1	467	807	10	0	100%	99%	0.779%	63%	100%

			True Negative - Fail OK	True Positive - Pass OK	False Negative - Yield Loss	False Positive - DPM	Fault Coverage - % of Correct Negative Pred - TN/TN+FP	Redundant Coverage - % of Correct Redundant Pred - TP/TP+FN	Accuracy - % of Correct Pred - TN+TP/ALL	Recall - % of Correct Positive - TP/TP+TN	Precision - % of Correct Positive Pred - TP/TP+FP	True Negative - Fail OK	True Positive - Pass OK	False Negative - Yield Loss	False Positive - DPM	Fault Coverage - % of Correct Negative Pred - TN/TN+FP	Redundant Coverage - % of Correct Redundant Pred - TP/TP+FN	Accuracy - % of Correct Pred - TN+TP/ALL	Recall - % of Correct Positive - TP/TP+TN	Precision - % of Correct Positive Pred - TP/TP+FP	
No Shift + Catast	Shift	0																			Process Parameter Variatio
	Process	400																			Catastrophic Faults +
	Catast	131																			Large Parametric Faults
	All	531																			
Shift + Catast w/o Etest	Shift	500																			Process Par + Catast + Large
	Process	400																			2 large process parameter s
	Catast	393																			No e-test data assumed
	All	1293																			
Shift + Catast w Etest	Shift	500																			Process Par + Catast + Large
	Process	400																			2 large process parameter s
	Catast	393																			E-test data used for stimulu
	All	1293																			classifier selection

Figure 104: Alternate SVM; detailed breakdown of experiments 2 and 3 with restricted training set in terms of evaluation metrics.

5.4.3 Dictionary-Independent Alternate Test with SVM

The third and the last method is to improve the go/no-go approach such that the final fault-based alternate test is defect-dictionary independent. In this scheme, only Monte-Carlo simulations of process variables are used for training, with no hard defects. The stimulus optimization is driven by an imaginary go/no-go boundary for the SVM, which is gradually progressed from the zero diameter centered at the optimal design to the diameter enclosing all Monte-Carlo simulations. The stimulus search is performed in the direction that best represents the progress in the parallel measurement space.

The motivation behind this methodology is to eliminate the main drawback of the Alternate SVM method: achieve similar figures of merit without the defect-dictionary. The supervised learner is again an SVM, implemented with WEKA. As any other SVM, it needs to be trained with both positive and negative examples to function properly. The solution presented here depends on virtual pass/fail boundaries. First, the specification space generated by S_{proc} is divided into co-centric hypercores. In order to simplify the explanation, let's think about a 2D mock-up specification space as in Figure 105(a) instead of the 7D specification space of our dataset. The instances in this toy space are generated by a Monte-Carlo simulation of process variations only, hence no access to a defect-dictionary. Each circle in this figure corresponds to a virtual classification boundary. Similarly, the corresponding measurement space is depicted in Figure 105(b) with respect to a given test stimulus. It is possible to train a series of SVMs such that first the instances lying inside the innermost circle are marked "pass" and the rest as "fail", then the instances lying inside the middle circle as "pass" and the rest as "fail", etc... These SVMs correspond to the virtual decision boundaries in Figure 105(b). As long as the topological similarity between the classification boundaries on the specification space and the decision boundaries on the measurement space is preserved over the entire range of classification boundaries,

the outermost decision boundary can be used as an educated guess for the actual SVM we have been looking for, the one between go/no-go instances. Consequently, the corresponding test stimulus can be assigned a high fitness value. As a negative example, consider Figure 105(c): the outermost decision boundary intersects with the decision boundary in the middle, hence it creates a contradiction in the progress towards the actual SVM; as far as we know -without the defect-dictionary- this input stimulus is not what we are looking for.

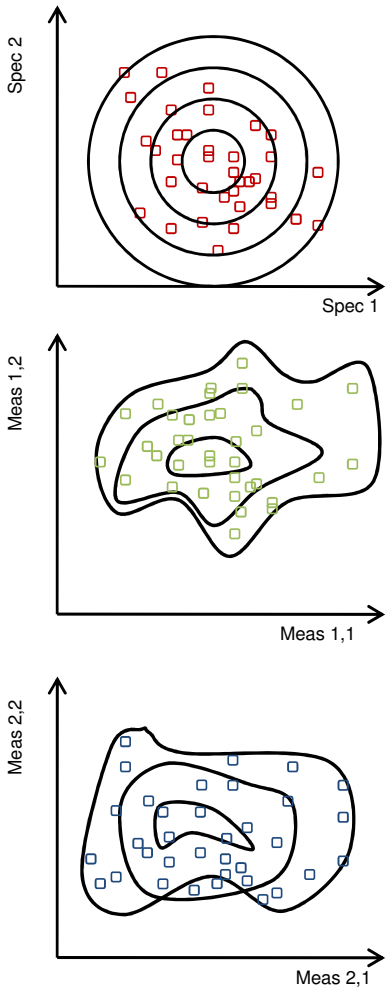


Figure 105: Simplified 2D virtual decision boundaries, (a) specification space, (b) measurement space 1, (c) measurement space 2.

The GA-based search algorithm uses a fitness function that measures topological

similarity and its consistency over successive virtual classification boundaries. The fitness value is an increasing function of the mean distance between two successive decision boundaries, which is measured using inner product spaces [140], and any double intersection points are heavily penalized. The GA search space, which are the circles in the 2D mockup, is populated by a brute-force algorithm, that seeds center points exhaustively through a grid defined by the specification resolution. Circles, or hypersurfaces in 7D, are created with uniform Δr increments until all samples are contained in a circle. The computational complexity of this universal population can be bounded by properly selecting Δr to be smaller than a fraction of the minimum distance between two instances, and larger than the specification resolution. In this case, complexity is $O(N^2)$, where N is the number of training instances. This is a one-time computational effort. For training sets of practical size, this computational effort is not significant with respect to the execution time of the GA.

Experiment 3 in the previous section is repeated with this classifier as well. The figures of merit for this method are listed in Table 5.4.3. When compared with the 6-bit/12-bit results are similar to alternate SVM. The FC of 98% for the 6-bit ADC is close to the acceptable range. Although PC is a little lower at 84%, this is again mostly due to the 1-to-100 penalty of FNs versus FPs. With e-test data, this method performs equivalent to alternate SVM, it is possible to achieve the figures of a 12-bit ADC with the 6-bit ADC (FC 100%, PC 99%, AC 99%). As expected, this method performs equally well without a defect-dictionary.

5.5 Conclusions

With a large dataset, 24 experiments are performed to explore four dimensions. The first dimension is the *choice of test methodology*, which is the main variable of the experiment. The properties of the three alternate test methodologies are compared along other dimensions of the experiment. *ATE complexity* constitutes the second

Table 17: Summary of Experiment 3 with dictionary-independent alternate SVM.

		Dictionary-Independent					
		12-bit ADC			**6-bit ADC**		
		FC	**PC**	**AC**	**FC**	**PC**	**AC**
Non-restricted	w/o Etest	100	92	95	98	84	89
training set	w/ Etest	100	99	99	100	99	99
Restricted	w/o Etest	100	92	95	98	84	89
training set	w/ Etest	100	99	99	100	99	99

dimension. The simulated responses are quantized with either 12 or 6 bits emulating effects of high or low-resolution ADCs. *Availability of e-test data* is the third dimension. When disabled, the alternate test methodology has one and only one optimized test stimulus for detecting faults. When enabled, they can make use of simple flags from wafer-probe testing and conditionally select from a set of previously optimized stimuli such that the selected stimulus best exploits the correlations in the subgroup designated by that e-test flag. The final dimension is the *availability of a defect-dictionary*. When the defect-dictionary is enabled, any instances from S_{def} can be used as part of the training sets; otherwise, a series of experiments are performed each prohibiting a certain type of defect from appearing in the training sets. This is especially critical for typical SVM-based methods, since they need those "negative" instances to develop robust go/no-go boundaries. Table 5.5 summarizes the results.

The three methodologies presented have different application profiles. The signature filter provides the most accurate solution given the resources. It is most suitable for third-party testing customers, who do not have access to e-test data and already invested in high-end ATEs. The alternate SVM provides flexibility in terms of ATE requirements; however, high yield can be achieved only with proper e-test data. Also, the process needs to be mature with enough hardware samples to account for a detailed defect-dictionary. The dictionary-independent SVM solution provides the same

benefits as the alternate SVM, even in low-volume production; however, the proposed test generation scheme is computationally more intensive. It requires an initial step to populate the GA space, computational complexity of this step scales with the second order of the number of training samples.

Table 18: Summary of the 24-step experiment.

		Signature Filter						**Alternate SVM**						**Dictionary Independent**					
		12-bit ADC			**6-bit ADC**			**12-bit ADC**			**6-bit ADC**			**12-bit ADC**			**6-bit ADC**		
		FC	**PC**	**AC**	**FC**	**PC**	**AC**	**FC**	**PC**	**AC**	**FC**	**PC**	**AC**	**FC**	**PC**	**AC**	**FC**	**PC**	**AC**
Non-restricted	w/o Etest	100	97	98	94	88	91	100	93	95	98	85	90	100	92	95	98	84	89
training set	w/ Etest	100	97	98	98	93	95	100	99	99	100	99	99	100	99	99	100	99	99
Restricted	w/o Etest	100	97	98	94	88	91	97	86	87	94	79	88	100	92	95	98	84	89
training set	w/ Etest	100	97	98	98	93	95	97	93	92	97	93	93	100	99	99	100	99	99

Chapter VI

CONCLUSIONS AND FUTURE WORK

Several enhancements to alternate testing is presented in this thesis. The methodologies presented enable low-cost RF testing for stand-alone and integrated components, but the main focus is on enabling efficient test of RF components integrated into SOPs.

Constrained-driven RF test stimulus generation presents techniques for tuning test generation process depending on available tester hardware, DfT features, and the amount of information available about the DUT. First example demonstrates a multisine stimulus generation with a simple greedy search and explores the tradeoffs between source complexity and alternate test prediction accuracy. Second example is focused on non-invasive BIT, and demonstrates the effectiveness of using an alternate test signature based on sampling of supply currents. Third and fourth examples explore the many dimensions of test generation including sample size, training algorithms, search algorithms, process variable distributions, and sampling noise. Specialized simulators and optimization algorithms are developed for test generation with behavioral models. A hardware experiment implements the technique on benchtop and ATE for a test case with IP constraints. Certain recipes are called out for combinations of test constraints. For complex systems such as transceivers, the speed of the proposed behavioral simulator can pose a bottleneck. If one considers the propagation of tones from the stimulus to the response, new tones are created through amplifiers and mixers while existing tones are crippled by filters and samplers. Hence, not all the tones generated by an earlier element in the chain show up in the response. In other words, the proposed simulator works with a "forward only" data flow model.

As an extension of the work presented in this thesis, an additional feature can be implemented by back propagating the filtering constraints from the last element to the first in the chain. At the beginning of the process the system is studied once, and valid frequency bands are assigned for the outputs of each element. Also for complex systems such as transceivers, the SUTs are usually end-to-end meaning that individual specifications of individual components like an LNA or a mixer are not critical. From this point of view, the proposed scheme presents additional benefits for diagnosis. The proposed end-to-end stimulus/response pair is used to predict individual specifications of the individual components as well as the cascade specifications. Note that for the diagnosis scenario to be effective, the complex system needs to have taps for corresponding component input/outputs so that relevant component specifications can be measured for training of the supervised learner.

Noise-referenced alternate testing presents a low-cost BIST scheme for RF components embedded in a system with DSP. The methodology is shown to predict complex specifications quantitatively by the use of alternate tests on spectral features generated by DSP algorithms. The BIST generation process is fully automated, hence proposes a designer-friendly DfT scheme. The scheme can also be extended to BIT of RF devices with low-cost digital testers. The methodology is verified by two examples that present close tracking of all the specifications even in the presence of imperfect stimuli. The case studies on input imperfections show that certain levels of noise, DC shift, phase shift and frequency shifts are tolerable. As an extension of the work presented in this thesis, selective instances can be added to the training set representing the expected imperfections. Normally, sweeping the range of imperfections quickly multiplies the size of training set, and makes test generation impractical. In this specific implementation, swept imperfections can be first evaluated with their corresponding frequency domain signature. Only those imperfections that significantly change the spectral signature need to be included in the training set. From this point

of view, noise-referenced alternate test shows the potential to implement a practical training set that can reflect severe input imperfections. This method cannot guarantee high prediction accuracy under severe imperfections, but can be used to define the quality of input given an accuracy target.

The recent literature on specification-based alternate test has shown that by using alternate testing the test specifications can be predicted very accurately, significantly reducing the cost. In order to use alternate test at frequencies in multi-GHz range, both the test waveforms need to be very simple and the evaluation of the test response should be handled by practical hardware-based test response feature extractors. These specialized analog circuits extract response signal waveform features in the form of low-bandwidth analog output signals. Furthermore, the built-in sensors used for measuring explicit figures of merits such as peaks, rms values, or zero crossings can be replaced with implicit signal feature extractors, which can make better use of the powerful mapping engine embedded in alternate test methodology. Case studies show that this scheme not only predicts many specifications accurately, but also serves as an auto-calibration capability for these sensors, which is validated in the case of temperature variations and large process shifts. As these sensors require little area in a BIT scheme, they can be implemented in between stages of an RF system, and the work can be expanded to diagnosis effectively.

The alternate testing scheme is also extended to defect-based domain. First, a simple signature filter is constructed by defining boundaries on the responses of a regular traning set. This filter acts as a preliminary step before the usual alternate mapping functions, and flags catastrophic faults. The filter is demonstrated to work effectively even in the presence of redundant faults. Second, the usual supervised learner, MARS, is replaced with a supervised classifier based on SVM. Regression is abandoned in favor of a go/no-go decision. Alternate testing with SVM is shown to reduce ATE complexity, at the expense of being dependent on a defect-dictionary.

The training set needs to reflect a fair distribution of expected faults. Third, a dictionary-independent SVM method is demonstrated. This scheme relies on negative samples, expected faults, defined by virtual pass/fail boundaries. By favoring SVMs that consistently perform well with changing pass/fail boundaries, over-training on a specific fault-dictionary is avoided. It is also shown that e-test data can be effectively used together with defect-based alternate test to guide stimulus generation and lower tester complexity. The work presented in this thesis can be expanded by implementing a hybrid scheme that uses a small defect-dictionary in addition to the dictionary-independent SVM. While populating the GA space, SVMs that perform poor on the small defect-dictionary can be automatically eliminated, which will reduce the amount of iterations for GA to converge on a solution.

REFERENCES

[1] ITRS Working Group, "International technology roadmap for semiconductors (ITRS), RF and A/MS technologies for wireless communications." `http://www.itrs.net/Links/2008ITRS/Home2008.htm`, September 12th, 2009.

[2] Vinnakota, B., ed., *Analog and Mixed-Signal Test*. Upper Saddle River, NJ: Prentice Hall, 1998.

[3] Burns, M. and Roberts, G. W., *An Introduction to Mixed-Signal IC Test and Measurement*. New York: Oxford Univ. Press, 2001.

[4] Schaub, K. and Kelly, J., *Production Testing of RF and System-on-a-Chip Devices for Wireless Communications*. Boston: Artech House, 2004.

[5] Tummala, R. R., *Fundamentals of Microsystems Packaging*. NY: McGraw-Hill, 2001.

[6] Yoon, H., Hou, J., Bhattacharya, S. K., Chatterjee, A., and Swaminathan, M., "Fault detection and automated fault diagnosis for embedded integrated electrical passives," *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol. 21, no. 3, pp. 265–276, 1999.

[7] Maly, W., Feltham, D., Gattiker, A., Hobaugh, M., Backus, K., and Thomas, M., "Smart-substrate multichip-module systems," *IEEE Design & Test of Computers*, vol. 11, no. 2, pp. 64–73, 1994.

[8] Gilg, L., "Known good die: a closer look," *Advanced Packaging*, vol. 14, pp. 24–7, Feb. 2005.

[9] Keezer, D., "Bare die testing and MCM probing techniques," in *Multi-Chip Module Conference, 1992. MCMC-92, Proceedings 1992 IEEE*, (Santa Cruz, CA), pp. 20–23, Mar. 1992.

[10] Berry, M., "How advances in RF and radio SiP affect test strategies," *Advanced Packaging*, vol. 14, pp. 21–4, Aug. 2005.

[11] Appello, D., Bernardi, P., Grosso, M., and Reorda, M., "System-in-package testing: problems and solutions," *IEEE Design & Test of Computers*, vol. 23, pp. 203–211, May/June 2006.

[12] Steps, S., "Full wafer test: making test more cost effective," in *Known-Good Die Workshop, 2005. Proceedings.*, 2005.

[13] Economikos, L., Morrison, T., and Crnic, F., "Electrical test of multichip substrates," *IEEE Transactions on Components, Packaging and Manufacturing Technology—Part B: Advanced Packaging*, vol. 17, pp. 56–61, Feb. 1994.

[14] Lee, K., Bhattacharya, S., Varadarajan, M., Wan, L., Abothu, I., Sundaram, V., Muthana, P., Balaraman, D., Raj, P., Swaminathan, M., Sitaraman, S., Tummala, R., Viswanadham, P., Dunford, S., and Lauffer, J., "Design, fabrication, and reliability assessment of embedded resistors and capacitors on multilayered organic substrates," *Advanced Packaging Materials: Processes, Properties and Interfaces, 2005. Proceedings. International Symposium on*, pp. 249–254, Mar. 2005.

[15] Ulrich, R. K. and Schaper, L. W., eds., *Integrated Passive Component Technology*. Wiley-IEEE Press, 2003.

[16] Johnson, A. J., "Efficient fault analysis in linear analog circuits," *IEEE Transactions on Circuits and Systems*, vol. 26, no. 7, pp. 475–484, 1979.

[17] Pan, C.-Y. and Cheng, K.-T., "Fault macromodeling for analog/mixed-signal circuits," in *Proc. International Test Conference*, (Washington, DC, USA), pp. 913–922, 1997.

[18] Milor, L. and Visvanathan, V., "Detection of catastrophic faults in analog integrated circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 8, pp. 114–130, Feb. 1989.

[19] Harvey, R., Richardson, A., Bruls, E., and Baker, K., "Analogue fault simulation based on layout dependent fault models," in *Proc. International Test Conference*, (Washington, DC, USA), pp. 641–649, 1994.

[20] Sebeke, C., Teixeira, J., and Ohletz, M., "Automatic fault extraction and simulation of layout realistic faults for integrated analogue circuits," in *Proc. European Design and Test Conference*, (Paris, France), pp. 464–468, 1995.

[21] Starzyk, J. and Dai, H., "Sensitivity based testing of nonlinear circuits," in *IEEE International Symposium on Circuits and Systems*, (Espoo, Finland), pp. 1159–1162, 1988.

[22] Hamida, N. B. and Kaminska, B., "Multiple fault analog circuit testing by sensitivity analysis," *Journal of Electronic Testing: Theory and Applications*, vol. 4, pp. 331–343, 1993.

[23] Michael, C. and Ismail, M., "Statistical modeling of device mismatch for analog MOS integrated circuits," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 2, pp. 154–166, 1992.

[24] Balivada, A., Zheng, H., Nagi, N., Chatterjee, A., and Abraham, J. A., "A unified approach for fault simulation of linear mixed-signal circuits," *Journal of Electronic Testing: Theory and Applications*, vol. 9, pp. 29–41, Aug. 1996.

[25] Nagi, N., Chatterjee, A., and Abraham, J. A., "Fault simulation of linear analog circuits," *Journal of Electronic Testing: Theory and Applications*, vol. 4, pp. 345–360, Nov. 1993.

[26] Sunter, S. and Nagi, N., "Test metrics for analog parametric faults," in *Proc. VLSI Test Symposium*, (Dana Point, CA, USA), pp. 226–234, 1999.

[27] Lindermeir, W., Graeb, H., and Antreich, K., "Design based analog testing by characteristic observation inference," in *Proc. International Conference on Computer-Aided Design*, (San Jose, CA, USA), pp. 620–626, 1995.

[28] Halder, A., Bhattacharya, S., and Chatterjee, A., "Automatic multitone alternate test-generaton for RF circuits using behavioral models," in *Proc. International Test Conference*, vol. 1, pp. 665–673, 2003.

[29] Sood, A., Choi, K., Haridass, A., Na, N., and Swaminathan, M., "Modeling and mixed signal simulation of embedded passive components in high performance packages," in *Proc. International Multichip Modules and High Density Packaging Conference*, (Denver, CO), pp. 506–511, Apr. 1998.

[30] Abadir, M., Parikh, A., Sandborn, P., Drake, K., and Bal, L., "Analyzing multichip module testing strategies," *IEEE Design & Test of Computers*, vol. 11, no. 1, pp. 40–52, 1994.

[31] Yoon, H., *Fault detection and identification techniques for embedded analog circuits*. Phd, Georgia Institute of Technology, Atlanta, GA, USA, 1998.

[32] Kornegay, K. and Roy, K., "Integrated test solutions and test economics for MCMs," in *Proc. International Test Conference*, (Washington, DC), pp. 193–201, Oct. 1995.

[33] Choi, K. L., *Modeling and simulation of embedded passives using rational functions in multi-layered substrates.* Phd, Georgia Institute of Technology, Atlanta, GA, USA, 1999.

[34] Kim, B. and Choi, H.-H., "A new test method for embedded passives in high density package substrates," in *Proc. Electronic Components and Technology Conference*, (Orlando, FL), pp. 1362–1366, May/June 2001.

[35] IEEE, "IEEE standard test access port and boundary-scan architecture." IEEE Std 1149.1-2001, 2001.

[36] Zorian, Y., "A structured testability approach for multi-chip modules based on BIST and boundary-scan," *IEEE Transactions on Advanced Packaging*, vol. 17, pp. 283–290, Aug. 1994.

[37] IEEE, "IEEE standard for a mixed-signal test bus." IEEE Std 1149.4-1999, 1999.

[38] Sasidhar, K., *Parallel test techniques for multi-chip modules.* Phd, Georgia Institute of Technology, Atlanta, GA, USA, 1997.

[39] Akbay, S. S., Halder, A., Chatterjee, A., and Keezer, D., "Low-cost test of embedded RF/analog/mixed-signal circuits in SOPs," *IEEE Transactions on Advanced Packaging*, vol. 27, no. 2, pp. 352–363, 2004.

[40] Keezer, D., Minier, D., and Caron, M., "A production-oriented multiplexing system for testing above 2.5Gbps," in *Proc. International Test Conference*, vol. 1, pp. 191–200, 2003.

[41] Davis, J. S., *An FPGA-based digital logic core for ATE support and embedded test applications.* Phd, Georgia Institute of Technology, Atlanta, GA, USA, 2003.

[42] Variyam, P., Cherubal, S., and Chatterjee, A., "Prediction of analog performance parameters using fast transient testing," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 3, pp. 349–361, 2002.

[43] Voorakaranam, R., Cherubal, S., and Chatterjee, A., "A signature test framework for rapid production testing of RF circuits," in *Proc. Design, Automation and Test in Europe*, (Paris, France), pp. 186–191, Mar. 2002.

[44] Bhattacharya, S. and Chatterjee, A., "Use of embedded sensors for built-in-test RF circuits," in *Proc. International Test Conference*, pp. 801–809, 2004.

[45] Akbay, S. S. and Chatterjee, A., "Feature extraction based built-in alternate test of RF components using a noise reference," in *Proc. VLSI Test Symposium*, pp. 273–278, 2004.

[46] Akbay, S. S. and Chatterjee, A., "Built-in test of RF components using mapped feature extraction sensors," in *Proc. VLSI Test Symposium*, pp. 243–248, 2005.

[47] Han, D., Akbay, S. S., Bhattacharya, S., Chatterjee, A., and Eisenstadt, W., "On-chip self-calibration of RF circuits using specification-driven built-in self-test (S-BIST)," in *Proc. International On-Line Testing Symposium*, pp. 106–111, 2005.

[48] Friedman, J. H., "Multivariate adaptive regression splines," *The Annals of Statistics*, vol. 19, pp. 1–141, Apr. 1991.

[49] Cadence Design Systems Inc., *Affirma RF Simulator (SpectreRF) User Guide*, 1999.

[50] Akbay, S. S. and Chatterjee, A., "Optimal multisine tests for RF amplifiers," in *Digest of Workshop on Test of Wireless Circuits and Systems*, (Baltimore, MD, USA), Oct. 2002.

[51] Devarayanadurg, G. and Soma, M., "Analytical fault modeling and static test generation for analog ICs," in *Proc. International Conference on Computer-Aided Design*, (San Jose, CA, USA), pp. 44–47, Nov. 1994.

[52] Nagi, N., Chatterjee, A., Balivada, A., and Abraham, J., "Fault-based automatic test generator for linear analog circuits," in *Proc. International Conference on Computer-Aided Design*, (Santa Clara, CA), pp. 88–91, 1993.

[53] Slamani, M. and Kaminska, B., "Multifrequency analysis of faults in analog circuits," *IEEE Design & Test of Computers*, vol. 12, no. 2, pp. 70–80, 1995.

[54] Tsai, S.-J., "Test vector generation for linear analog devices," in *Proc. International Test Conference*, Oct. 1991.

[55] Huynh, S., Kim, S., Soma, M., and Zhang, J., "Automatic analog test signal generation using multi frequency analysis," *IEEE Transactions on Circuits and Systems—Part II: Analog and Digital Signal Processing*, vol. 46, no. 5, pp. 565–576, 1999.

[56] Variyam, P. N., Cherubal, S., and Chatterjee, A., "Specification driven test design for analog circuits." U.S. Patent application serial no. 09/575488.

[57] Dufort, B. and Roberts, G., "On-chip analog signal generation for mixed-signal built-in self-test," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 3, pp. 318–330, 1999.

[58] Kim, S. and Soma, M., "An all-digital built-in self-test for high-speed phase-locked loops," *IEEE Transactions on Circuits and Systems—Part II: Analog and Digital Signal Processing*, vol. 48, no. 2, pp. 141–150, 2001.

[59] Yoon, J.-S. and Eisenstadt, W., "Embedded loopback test for RF ICs," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 5, pp. 1715–1720, 2005.

[60] Hsieh, H.-H. and Lu, L.-H., "Integrated CMOS power sensors for RF BIST applications," in *Proc. VLSI Test Symposium*, Apr. 2006.

[61] Wang, Q. and Soma, M., "RF front-end system gain and linearity built-in test," in *Proc. VLSI Test Symposium*, Apr. 2006.

[62] Valdes-Garcia, A., Silva-Martinez, J., and Sanchez-Sinencio, E., "On-chip testing techniques for RF wireless transceivers," *IEEE Design & Test of Computers*, vol. 23, no. 4, pp. 268–277, 2006.

[63] Cherubal, S. and Chatterjee, A., "Test generation based diagnosis of device parameters for analog circuits," in *Proc. Design, Automation and Test in Europe*, (Munich, Germany), pp. 596–602, 2001.

[64] Variyam, P. and Chatterjee, A., "Digital-compatible BIST for analog circuits using transient response sampling," *IEEE Design & Test of Computers*, vol. 17, no. 3, pp. 106–115, 2000.

[65] Halder, A. and Chatterjee, A., "Specification-based digital compatible built-in test of embedded analog circuits," in *Proc. Asian Test Symposium*, (Kyoto, Japan), pp. 344–349, 2001.

[66] Bhattacharya, S. and Chatterjee, A., "A DFT approach for testing embedded systems using DC sensors," *IEEE Design & Test of Computers*, vol. 23, no. 6, pp. 464–475, 2006.

[67] Yin, Q., Eisenstadt, W., and Fox, R., "A translinear-based RF RMS detector for embedded test," in *Proc. International Symposium on Circuits and Systems*, vol. 1, 2004.

[68] Kim, H., Walker, D., and Colby, D., "A practical built-in current sensor for IDDQ testing," in *Proc. International Test Conference*, (Baltimore, MD), pp. 405–414, 2001.

[69] Miller, A., "IDDQ testing in deep submicron integrated circuits," in *Proc. International Test Conference*, (Atlantic City, NJ), pp. 724–729, 1999.

[70] Gattiker, A. and Maly, W., "Current signatures [VLSI circuit testing]," in *Proc. VLSI Test Symposium*, (Princeton, NJ), pp. 112–117, 1996.

[71] Dragic, S. and Margala, M., "A 1.2V built-in architecture for high frequency on-line IDDQ/delta IDDQ test," in *Proc. Computer Society Annual Symposium on VLSI*, (Pittsburgh, PA), pp. 148–153, 2002.

[72] Sidiropulos, M., Stopjakova, V., and Manhaeve, H., "Implementation of a BIC monitor in a new analog BIST structure," in *IDDQ Testing, 1996., IEEE International Workshop on*, (Washington, DC), pp. 59–63, 1996.

[73] Srinivasan, G., Bhattacharya, S., Cherubal, S., and Chatterjee, A., "Efficient test strategy for TDMA power amplifiers using transient current measurements: uses and benefits," in *Proc. Design, Automation and Test in Europe*, vol. 1, pp. 280–285, 2004.

[74] Gopalan, A., Das, T., Washburn, C., and Mukund, P., "An ultra-fast, on-chip BiST for RF low noise amplifiers," in *Proc. International VLSI Design Conference*, pp. 485–490, Jan. 2005.

[75] Soldo, A., Gopalan, A., Mukund, P., and Margala, M., "A current sensor for on-chip, non-intrusive testing of RF systems," in *Proc. International VLSI Design Conference*, pp. 1023–1026, Jan. 2004.

[76] Bhattacharya, S., Srinivasan, G., Cherubal, S., Halder, A., and Chatterjee, A., "System-level testing of RF transmitter specifications using optimized periodic bitstreams," in *Proc. VLSI Test Symposium*, pp. 229–234, 2004.

[77] Halder, A. and Chatterjee, A., "Automated test generation and test point selection for specification test of analog circuits," in *Proc. International Symposium on Quality Electronic Design*, pp. 401–406, 2004.

[78] Voorakaranam, R. and Chatterjee, A., "Test generation for accurate prediction of analog specifications," in *Proc. VLSI Test Symposium*, (Montreal, Que., Canada), pp. 137–142, Apr. 2000.

[79] Voorakaranam, R., Akbay, S. S., Bhattacharya, S., Cherubal, S., and Chatterjee, A., "Signature testing of analog and RF circuits: algorithms and methodology," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 54, pp. 1018–1023, May 2007.

[80] Bracken, J., Raghavan, V., and Rohrer, R., "Extension of the asymptotic waveform evaluation technique with the method of characteristics," in *Proc. International Conference on Computer-Aided Design*, (Santa Clara, CA), pp. 71–75, 1992.

[81] Wood, J., Qin, X., and Cognata, A., "Nonlinear microwave/RF system design and simulation using agilent ADS system - data models," in *Proc. IEEE International Workshop on Behavioral Modeling and Simulation*, pp. 75–79, 2002.

[82] Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R., "Least angle regression," *Annals of Statistics (with discussion)*, vol. 32, no. 2, pp. 407–499, 2004.

[83] Friedman, J. H., "Function approximation: a gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–232, 2001.

[84] Keerthi, S., Shevade, S., Bhattacharyya, C., and Murthy, K., "Improvements to Platt's SMO algorithm for SVM classifier design," *Neural Computation*, vol. 13, no. 3, pp. 637–649, 2001.

[85] Witten, I. H. and Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco: Morgan Kaufmann, 2 ed., 2005.

[86] Vafaie, H. and Imam, I., "Feature selection methods: Genetic algorithms vs greedy-like search," in *Proc. International Conference on Fuzzy and Intelligent Control Systems*, 1994.

[87] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*. Massachusetts: Addison-Wesley, 1989.

[88] Wu, C.-Y. and Kao, H.-S., "A 2-V low-power CMOS direct-conversion quadrature modulator withintegrated quadrature voltage-controlled oscillator and RF amplifier forGHz RF transmitter applications," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 49, pp. 123–134, Feb. 2002.

[89] Zheng, R., Li, W., Li, N., and Ren, J., "An RF transmitter front-end implementation in 0.18m CMOS for OFDM-UWB," in *Solid-State and Integrated-Circuit Technology, 2008. ICSICT 2008. 9th International Conference on*, (Beijing,), pp. 1564–1568, Oct. 2008.

[90] RFMD, "Dual band front-end RF2411, LNA & mixer datasheet, RF Micro Devices." `http://www.rfmd.com`, October 16th, 2004.

[91] Consortium, S. T., *Openstar ATE System Platform Specifications*. Semiconductor Test Consortium, Nov. 2004.

[92] Agilent, *Signal Studio for Enhanced Multitone Manual*. Agilent.

[93] Hastie, T., Tibshirani, R., and Friedman, J. H., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York, NY: Springer Series in Statistics, 2001.

[94] Friedman, J. H., "An overview of computational learning and function approximation," in *From Statistics to Neural Networks. Theory and Pattern Recognition Applications* (Cherkassy, Friedman, and Wechsler, eds.), Springer-Verlag, 1994.

[95] Chatterjee, A., Kim, B., and Nagi, N., "DC built-in self-test for linear analog circuits," *IEEE Design & Test of Computers*, vol. 13, no. 2, pp. 26–33, 1996.

[96] Wey, C.-L., "Built-in self-test (BIST) structure for analog circuit fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 39, no. 3, pp. 517–521, 1990.

[97] Toner, M. and Roberts, G., "A BIST scheme for a SNR, gain tracking, and frequency response testof a sigma-delta ADC," *IEEE Transactions on Circuits*

and Systems—Part II: Analog and Digital Signal Processing, vol. 42, no. 1, pp. 1–15, 1995.

[98] Ohletz, M., "Hybrid built-in self-test (HBIST) for mixed analogue/digital integrated circuits," in *Proc. European Test Conference*, (Munich, Ger), Apr. 1991.

[99] Singhal, K. and Pinel, J., "Statistical design centering and tolerancing using parametric sampling," *IEEE Transactions on Circuits and Systems*, vol. 28, no. 7, pp. 692–702, 1981.

[100] Hemink, G., Meijer, B., and Kerkhoff, H., "Testability analysis of analog systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, no. 6, pp. 573–583, 1990.

[101] Wagner, K. and Williams, T., "Design for testability of analog/digital networks," *IEEE Transactions on Industrial Electronics*, vol. 36, no. 2, pp. 227–230, 1989.

[102] Slamani, M. and Kaminska, B., "Fault observability analysis of analog circuits in frequency domain," *IEEE Transactions on Circuits and Systems—Part II: Analog and Digital Signal Processing*, vol. 43, no. 2, pp. 134–139, 1996.

[103] Stenbakken, G. N. and Souders, T. M., "Test-point selection and testability measures via QR factorization of linear models," *IEEE Transactions on Instrumentation and Measurement*, vol. 36, pp. 406–410, June 1986.

[104] Fasang, P., "Boundary scan and its application to analog-digital ASIC testing in a board/system environment," in *Proc. Custom Integrated Circuits Conference*, (San Diego, CA, USA), pp. 1–22, 1989.

[105] Sunter, S., "The p1149.4 mixed signal test bus: costs and benefits," in *Proc. International Test Conference*, (Washington, DC, USA), pp. 444–450, 1995.

[106] Duhamel, P. and Rault, J., "Automatic test generation techniques for analog circuits and systems: A review," *IEEE Transactions on Circuits and Systems*, vol. 26, no. 7, pp. 411–440, 1979.

[107] Sunter, S. and Roy, A., "BIST for phase-locked loops in digital applications," in *Proc. International Test Conference*, (Atlantic City, NJ, USA), pp. 532–540, 1999.

[108] Kim, S. and Soma, M., "Test evaluation and data on defect-oriented BIST architecture for high-speed PLL," in *Proc. International Test Conference*, (Baltimore, MD, USA), pp. 830–837, 2001.

[109] Hafed, M., Abaskharoun, N., and Roberts, G., "A 4-GHz effective sample rate integrated test core for analog and mixed-signal circuits," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 4, pp. 499–514, 2002.

[110] Mendez-Rivera, M., Silva-Martinez, J., and Sanchez-Sinencio, E., "On-chip spectrum analyzer for built-in testing analog ICs," in *Proc. International Symposium on Circuits and Systems*, vol. 5, 2002.

[111] Hawrysh, E. and Roberts, G., "An integration of memory-based analog signal generation into current DFT architectures," in *Proc. International Test Conference*, (Washington, DC), pp. 528–537, 1996.

[112] Kao, W. and Xia, J., "Automatic synthesis of DUT board circuits for testing of mixed signal ICs," in *Proc. VLSI Test Symposium*, (Atlantic City, NJ, USA), pp. 230–236, 1993.

[113] Huertas, G., Vazquez, D., Peralias, E., Rueda, A., and Huertas, J., "Practical oscillation-based test of integrated filters," *IEEE Design & Test of Computers*, vol. 19, no. 6, pp. 64–72, 2002.

[114] Arabi, K. and Kaminska, B., "Oscillation-test methodology for low-cost testing of active analog filters," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 4, pp. 798–806, 1999.

[115] Arabi, K. and Kaminska, B., "Testing analog and mixed-signal integrated circuits using oscillation-test method," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 16, no. 7, pp. 745–753, 1997.

[116] Raghunathan, A., Shin, H. J., and Abraham, J., "Prediction of analog performance parameters using oscillation based test," in *Proc. VLSI Test Symposium*, pp. 377–382, 2004.

[117] Negreiros, M., Carro, L., and Susin, A., "A statistical sampler for a new on-line analog test method," in *Proc. International On-Line Testing Symposium*, pp. 79–83, 2002.

[118] Negreiros, M., Carro, L., and Susin, A., "Ultra low cost analog BIST using spectral analysis," in *Proc. VLSI Test Symposium*, pp. 77–82, 2003.

[119] Nag, P., Gattiker, A., Wei, S., Blanton, R., and Maly, W., "Modeling the economics of testing: a DFT perspective," *IEEE Design & Test of Computers*, vol. 19, no. 1, pp. 29–41, 2002.

[120] Williams, D. and Ambler, A., "System manufacturing test cost model," in *Proc. International Test Conference*, pp. 482–490, 2002.

[121] Turino, J., "Test economics in the 2lst century," *IEEE Design & Test of Computers*, vol. 14, no. 3, pp. 41–44, 1997.

[122] Zorian, Y., "Testing the monster chip," *IEEE Spectrum*, vol. 36, no. 7, pp. 54–60, 1999.

[123] ITRS Working Group, "International technology roadmap for semiconductors (ITRS), test and test equipment." `http://www.itrs.net/Links/2008ITRS/Home2008.htm`, September 12th, 2009.

[124] Negreiros, M., Carro, L., and Susin, A., "Ultimate low cost analog BIST," in *Proc. Design Automation Conference*, pp. 570–573, 2003.

[125] Papoulis, A., *Probability, Random Variables, and Stochastic Processes*. NY: McGraw-Hill, 3rd ed., 1991.

[126] Bendat, J. and Piersol, A., *Random Data: Analysis and Measurement Procedures*. Wiley, 1986.

[127] Daubechies, I., "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[128] Donoho, D., "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[129] Zhang, T., Eisenstadt, W., and Fox, R., "Embedded self test for a 5GHz LNA," in *GSRC/SRC Workshop on MS, RF and Multi-GHz Testing*, (Napa Valley), Apr. 2004.

[130] Bhattacharya, S. and Chatterjee, A., "Built-in-test of analog and RF circuits using embedded sensors," in *Test Resource Partitioning Workshop*, (Napa Valley), Apr. 2004. `http://www.ece.gatech.edu/research/labs/mcmtest/`, October 21st, 2007.

[131] Meyer, R., "Low-power monolithic RF peak detector analysis," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 1, pp. 65–67, 1995.

[132] Zhang, T., Eisenstadt, W., and Fox, R., "A novel 5GHz RF power detector," in *Proc. International Symposium on Circuits and Systems*, vol. 1, pp. 897–900, 2004.

[133] McKeon, A. and Wakeling, A., "Fault diagnosis in analogue circuits using AI techniques," in *Proc. International Test Conference*, (Washington, DC, USA), pp. 118–123, Aug. 1989.

[134] Voorakaranam, R., Akbay, S. S., Bhattacharya, S., Cherubal, S., and Chatterjee, A., "Signature testing of analog and RF circuits: Algorithms and methodology," *IEEE Transactions on Circuits and Systems*, vol. 54, pp. 1018–1031, May 2007.

[135] Vapnik, V., *Statistical Learning Theory*. Wiley, 1998.

[136] Stratigopoulos, H.-G. and Makris, Y., "Nonlinear decision boundaries for testing analog circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 11, pp. 1760–1773, 2005.

[137] Stratigopoulos, H.-G. and Makris, Y., "Bridging the accuracy of functional and machine-learning-based mixed-signal testing," in *Proc. VLSI Test Symposium*, Apr./May 2006.

[138] Scholkopf, B. and Smola, A., eds., *Leaning with Kernels*. MIT Press, 2001.

[139] Witten, I. H. and Frank, E., *Data Mining: Practical machine learning tools and techniques*. San Francisco: Morgan Kaufmann, 2005.

[140] Wahba, G., "An introduction to model building with reproducing kernel hilbert spaces," Tech. Rep. 1020, DEPARTMENT OF STATISTICS, University

of Wisconsin, 1210 West Dayton St., Madison, WI 53706, April 2000. Short course notes.

VITA

Selim Sermet Akbay was born on June 11th, 1977 in Ankara, Turkey. In 2000, he received the B.S. degree in Electrical and Electronics Engineering, from the prestigious Middle East Technical University, Ankara, Turkey. He was also accepted to the double major program and received his second B.S. degree in Computer Engineering. During his undergrad, he worked as a part-time research assistant in MEMS/VLSI Design Lab under the supervision of Professor Tayfun Akin. He joined School of Electrical and Computer Engineering at Georgia Institute of Technology in August 2000 for his graduate studies where he completed his M.S. and PhD in 2002 and 2009 respectively.

He was a graduate research assistant in the Testing and Reliability Engineering Lab under the supervision of Professor Abhijit Chatterjee. His research interests were the test of analog and radio-frequency electronic circuits, test automation and built-in self-test. He is currently with Intel Corporation.

Sermet received the Colonel Oscar P. Cleaver Award, Outstanding Graduate Student in the School of Electrical and Computer Engineering, Georgia Institute of Technology in 2000. His paper, which documents the results of a 6-month coop in Intel Corporation, received the Best in DTTC Award for the Intel DTTC Conference 2006.